



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C07K 14/435, 14/47, C12N 1/21, 5/00,</b> <b>15/12, 15/63</b>	<b>A1</b>	<b>(11) International Publication Number:      WO 99/46289</b>  <b>(43) International Publication Date:    16 September 1999 (16.09.99)</b>
<b>(21) International Application Number:      PCT/US99/05721</b>  <b>(22) International Filing Date:              11 March 1999 (11.03.99)</b>  <b>(30) Priority Data:</b> 60/077,686              12 March 1998 (12.03.98)      US 60/077,687              12 March 1998 (12.03.98)      US 60/077,696              12 March 1998 (12.03.98)      US 60/077,714              12 March 1998 (12.03.98)      US  <b>(71) Applicant (for all designated States except US):</b> HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). FERRIE, Ann, M. [US/US]; 120 Fox Run Drive, Tewksbury, MA 01876 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). FLORENCE, Charles [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). YU, Guo-Liang [CN/US]; 242 Gravatt Drive, Berkeley, CA 94705 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US).	<b>(74) Agents:</b> BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).  <b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>With an indication in relation to deposited biological</i> <i>material furnished under Rule 13bis separately from the</i> <i>description.</i>	
<b>(54) Title: 31 HUMAN SECRETED PROTEINS</b>  <b>(57) Abstract</b>  <p>The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## 31 Human Secreted Proteins

### *Field of the Invention*

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and  
5 their production.

### *Background of the Invention*

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or  
10 organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum  
15 (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or  
25 secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include  
30 the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using  
35 secreted proteins or the genes that encode them.

### *Summary of the Invention*

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

### *Detailed Description*

#### Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

In specific embodiments, the polynucleotides of the invention are less than 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, or 7.5 kb in length. In a further embodiment, polynucleotides of the invention comprise at least 15 contiguous nucleotides of the coding sequence, but do not comprise all or a portion of any intron. In another embodiment, the nucleic acid comprising the coding sequence does not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene in the genome).

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence



of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence.

Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42°C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH<sub>2</sub>PO<sub>4</sub>; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress

background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxiribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or

without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

### **Polynucleotides and Polypeptides of the Invention**

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 1**

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: NYFPVHTVQPNWYV (SEQ ID NO:77). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding

the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

5 This gene is expressed primarily in whole brain and infant brain tissues, and to a lesser extent in T-cells and fetal lung tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative disorders. Similarly, polypeptides and antibodies  
10 directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g.,  
15 lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in neural tissues such as infant and whole brain tissues  
20 indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of neurodegenerative disorders. Furthermore, the tissue distribution in brain tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons  
25 Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception.

In addition, the gene or gene product may also play a role in the treatment and/or  
30 detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The first approximately 333 nt of sequence shown in the sequence listing is vector sequence which will immediately be recognized by those of skill in the art.

35 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present

invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 776 of SEQ ID NO:11, b is an integer of 15 to 790, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

## FEATURES OF PROTEIN ENCODED BY GENE NO: 2

10 This gene is expressed primarily in colon tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, gastrointestinal disorders and colon cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the gastrointestinal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., gastrointestinal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 45 as residues: Ser-69 to Lys-74.

The tissue distribution in colon tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of colon cancer. Furthermore, the tissue distribution in gastrointestinal tissues (colon) indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, prevention, and/or treatment of various metabolic disorders such as Tay-Sachs disease, phenylketonuria, galactosemia, porphyrias, and Hurler's syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the

scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 540 of SEQ ID NO:12, b is an integer of 15 to 554, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where b is greater than or equal to a + 14.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 3

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: PVFTVNFLAWVHAPPVSITVDLIPTLAQAWS (SEQ ID NO:78). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in colon tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, gastrointestinal disorders and colon cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the gastrointestinal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., gastrointestinal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in colon tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of colon cancer. Furthermore, the tissue distribution in gastrointestinal tissues (colon) indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, prevention, and/or treatment of various metabolic disorders such as Tay-Sachs disease, phenylketonuria, galactosemia, porphyrias, and Hurler's syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present

invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1092 of SEQ ID NO:13, b is an integer of 15 to 1106, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 4**

10 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: WIQRIRTSADQLGPKKVVXFGLACCGVSGLFYA (SEQ ID NO:79). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in CD34 positive cells.

15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation, allergy and graft rejection, and immune system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in  
20 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,  
25 synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in CD34 positive cells indicates that polynucleotides and  
30 polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of hematopoietic and immune disorders such as inflammation, as well as immune modulation and differentiation. Furthermore, expression of this gene product in CD34 positive cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including  
35 blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Those of skill in the art will recognize that some vector nucleotide sequence is contained at the 5' and 3' ends of the sequence shown for this gene in the sequence listing.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 554 of SEQ ID NO:14, b is an integer of 15 to 568, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 5

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: PPGLCAAIP LQTRSAQGPWGGRRQSGGWCWGT VVGS GSS GGGNAFTGLGPVSTLPSLHGKQGVTSVTCHGGYVYTTGRXGAYYQLFVRDG QLQPVLRQKSCRGMNWL AGLRIVPDGSMVILGFHANEFVVWNPR SHEKLHIV NCGGGHRSWAFSDTEAAMAFAYLKDGDVMLYRALGGCTRPHVILREGLHGR EITCVKRVGTITLGPEYGVPSFMQPDDLEPGSEGPD LTDIVITCSEDTTVCVLALP TTTGSAHALTAVCNHISSVRAVAVWGIGTPGGPQDPQPGLTAHVVSAGGRAE MHCFSIMVTPDPSTPSRLACHVMHLXSHRLDEYWDRQRNRHRMVKVDPETR (SEQ ID NO:80), PPGLCAAIP LQTRSAQGPWGGRRQSGGWCWGT VVGS GSS (SEQ ID NO:81), GGGNAFTGLGPVSTLPSLHGKQGVTSVTCHGGYVYTTGRX (SEQ ID NO:82), GAYYQLFVRDGQLQPVLRQKSCRGMNWL AGLRIVPDGSMV (SEQ ID NO:83), ILGFHANEFVVWNPR SHEKLHIVNCGGGHRSWAFSDTEAAM (SEQ ID NO:84), AFAYLKDGDVMLYRALGGCTRPHVILREGLHGREITCVKRV



G (SEQ ID NO:85), TTLGPEYGVPSFMQPDDLEPGSEGPLDITDIVITCSEDTTVVCV (SEQ ID NO:86), LALPTTTGSAHALTAVCNHISSVRAVAVWGIGTPGGPQDPQ (SEQ ID NO:87), PGLTAHVVSAGGRAEMHCFSIMVTPDPSTPSRLACHVMHL (SEQ ID NO:88), and/or XSHRLDEYWDRQRNRHRMVKVDPETR (SEQ ID NO:89). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in LNCAP untreated cell line and endometrial tumor tissue, and to a lesser extent in other cancerous tissues such as adrenal gland tumor tissues and synovial sarcoma tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers, i.e., uncontrolled cell proliferation and/or differentiation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate and endometrial tissues, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., reproductive, gastrointestinal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 48 as residues: Lys-37 to Ile-45.

The tissue distribution in cancerous tissues, such as cancerous tissues of the endometrium, synovium, and adrenal gland tissues, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of tumors, as well as for regulating cell proliferation and/or differentiation. Expression within cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ

ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more  
 5 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3678 of SEQ ID NO:15, b is an integer of 15 to 3692, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a + 14.

## 10 FEATURES OF PROTEIN ENCODED BY GENE NO: 6

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: LMSLLTSPHQPPPPPPASASPSAVPNGPQSPKQQKEPL SHRFNEFMTSKPKIHCFRSLKRGVSSAPESCLSGVLWLHVWFCTITNFVCE (SEQ ID NO:90); FQNAKEEASVLPYVETVFLFGGGIFAMALCLISDALSSYR  
 15 DSHTNRVLTSPPF (SEQ ID NO:91); and/or RLMPFPSPSPRLLVTLAGREDVV GHSCNTLSAHLLEIVTMLITWF (SEQ ID NO:92). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is thought to reside on chromosome 9. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 9.

20 This gene is expressed primarily in activated T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders. Similarly, polypeptides and antibodies directed to  
 25 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample  
 30 taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution primarily in T-cells indicates that polynucleotides and  
 35 polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of immune disorders involving activated T-cells, e.g., in diseases relating to improper thymus, liver, and/or spleen function. Furthermore, expression of this gene product in

T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in T cells also strongly indicates a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1414 of SEQ ID NO:16, b is an integer of 15 to 1428, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 7**

The translation product of this gene shares sequence homology with a rat potassium-dependent sodium-calcium exchanger (See Genbank Accession No. gil2662461), as well as one from Bos taurus. These proteins are thought to be important in modulating Ca<sup>2+</sup> flux across the rod outer segments (ROS) of the retinal rod photoreceptors.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: GGXDDDEGPYTPFDTPSGKLETVKWAFTWPLSFVLYF

TVPNCNKPRWEKWF (SEQ ID NO:93). Polynucleotides encoding these polypeptides are also encompassed by the invention.

When tested against Jurkat cell lines, supernatants removed from cells containing this gene activated the NF-kB transcription factor. Thus, it is likely that this gene activates Jurkat cells, and to a lesser extent other immune cells, by activating a transcriptional factor found within these cells. Nuclear factor kB is a transcription factor activated by a wide variety of agents, leading to cell activation, differentiation, or apoptosis. Reporter constructs utilizing the NF-kB promoter element are used to screen supernatants for such activity.

Additionally, when tested against Jurkat cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates Jurkat cells, and to a lesser extent in other immune cells, through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

Likewise, when tested against K562 leukemia cell lines, supernatants removed from cells containing this gene activated the ISRE assay. Thus, it is likely that this gene activates leukemia cells, and to a lesser extent other cells, through the Jak-STAT signal transduction pathway. The interferon-sensitive response element is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in fetal and infant brain tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, color blindness, light sensitivity and neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the optic and neurological systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., optic, neural, cancerous

and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5       The tissue distribution in fetal and infant brain tissues, and the homology to retinal potassium-dependent sodium-calcium exchanger gene, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of various optic disorders related to light adaptation in rod  
10       distribution in brain tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS,  
15       psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception.

      In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein  
20       may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

      Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present  
25       invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1475 of SEQ ID NO:17, b is an integer of 15  
30       to 1489, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 8**

      The gene encoding the disclosed cDNA is thought to reside on chromosome 17.  
35       Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

This gene is expressed primarily in placental tissue, and to a lesser extent in breast tissue and melanocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast cancer and melanoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, metabolic and integumental systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, metabolic, integumentary, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in placental and breast tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of certain cancers, including breast cancer and melanomas. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1926 of SEQ ID NO:18, b is an integer of 15 to 1940, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 9**

The translation product of this gene shares a very small block of sequence  
5 homology with human hemopoietic cell protein-tyrosine kinase (HCK). The hck gene  
encodes a 505-residue polypeptide that is closely related to pp56lck, a lymphocyte-  
specific protein-tyrosine kinase. The exon breakpoints of the hck gene, partially defined  
by using murine genomic genes, demonstrate that hck is a member of the src gene  
family and has been subjected to strong selection pressure during mammalian evolution.  
10 High-level expression of hck transcripts in granulocytes is especially provocative since  
these cells are terminally differentiated and typically survive in vivo for only a few  
hours.

Thus the hck gene, like other members of the src gene family, appears to  
function primarily in cells with little growth potential. The translation product of this  
15 gene is expected to share certain biological activities with HCK based on the sequence  
similarity between the proteins. The gene encoding the disclosed cDNA is thought to  
reside on chromosome 20. Accordingly, polynucleotides related to this invention are  
useful as a marker in linkage analysis for chromosome 20.

This gene is expressed primarily in human prostate cancer, and to a lesser extent  
20 in activated neutrophils and primary dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as  
reagents for differential identification of the tissue(s) or cell type(s) present in a  
biological sample and for diagnosis of diseases and conditions which include, but are  
not limited to, prostate cancer; hematopoietic disorders; immune dysfunction;  
25 susceptibility to infection; and inflammation. Similarly, polypeptides and antibodies  
directed to these polypeptides are useful in providing immunological probes for  
differential identification of the tissue(s) or cell type(s). For a number of disorders of  
the above tissues or cells, particularly of the prostate and/or immune system, expression  
of this gene at significantly higher or lower levels may be routinely detected in certain  
30 tissues or cell types (e.g., gastrointestinal, immune, cancerous and wounded tissues) or  
bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or  
another tissue or cell sample taken from an individual having such a disorder, relative to  
the standard gene expression level, i.e., the expression level in healthy tissue or bodily  
fluid from an individual not having the disorder.

35 The tissue distribution in prostate cancer tissue, dendritic cells and neutrophils,  
and the short block of homology to hck, indicates that polynucleotides and polypeptides  
corresponding to this gene are useful for the diagnosis and/or treatment of prostate

cancer, as well as disorders of the immune system. For example, this gene product is thought to play a role in the abnormal cellular proliferation that accompanies prostate cancer. Inhibitors of the action of this gene product have beneficial therapeutic application in the treatment of prostate cancer. Alternately, the expression in neutrophils and dendritic cells indicates that this gene product may play a role in the survival, proliferation, and/or differentiation of hematopoietic cells, and may play key roles in inflammation and immunity. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1578 of SEQ ID NO:19, b is an integer of 15 to 1592, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 10**

The gene encoding the disclosed cDNA is thought to reside on chromosome 13. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 13.

This gene is expressed primarily in primary dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders; defects in immunity; susceptibility to infections; and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the



expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 53 as residues: Glu-35 to Lys-44, Cys-83 to Gly-88.

5       The tissue distribution in primary dendritic cells indicates that protein products of this gene are useful for the diagnosis and/or treatment of a variety of immune disorders. Expression of this gene product by dendritic cells indicates that it may play a role in the immune recognition/presentation process and may therefore be involved in the regulation of immunity. Alternately, it may represent a protein that is produced by  
10       cells, such as dendritic cells, that then has an effect on other hematopoietic cell types, and thereby regulates their survival, proliferation, activation, and/or differentiation. Therefore this gene product may have therapeutic benefit in a variety of hematopoietic disorders. Furthermore, expression of this gene product in primary dendritic cells also strongly indicates a role for this protein in immune function and immune surveillance.  
15       Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present  
20       invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1396 of SEQ ID NO:20, b is an integer of 15  
25       to 1410, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 11**

This gene is expressed primarily in primary dendritic cells.

30       Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic disorders; immune dysfunction; impaired immunity; and susceptibility to infections. Similarly, polypeptides and antibodies directed to these  
35       polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or

lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 54 as residues: Ala-107 to Ser-112.

The tissue distribution in primary dendritic cells indicates that protein products of this gene are useful for the diagnosis and/or treatment of hematopoietic disorders. Expression of this gene product specifically in primary dendritic cells indicates that it may play a role in immune responses. Therefore, it may have clinical utility in a variety of disorders that are characterized by impaired immune function or susceptibility to infection. Alternately, it may represent a gene product that is produced by specific cells, such as dendritic cells, that has effects on the survival, activation, proliferation, and/or differentiation of other cell types, most notably other hematopoietic cell lineages. Therefore, the gene product may have clinical utility in the treatment of a variety of hematopoietic disorders. Furthermore, expression of this gene product in primary dendritic cells also strongly indicates a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1713 of SEQ ID NO:21, b is an integer of 15 to 1727, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 12**

This gene is expressed primarily in primary dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic disorders; immune dysfunction; susceptibility to infection; and inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 55 as residues: Ser-106 to Leu-113.

The tissue distribution in primary dendritic cells indicates that the protein products of this gene are useful for the diagnosis and/or treatment of immune disorders. Expression of this gene specifically in dendritic cells indicates a role in immune function. Therefore, this gene product may be clinically useful in disorders marked by impaired or altered immune function, such as susceptibility to infection or impaired immune surveillance. Alternately, this may represent a gene product that is produced by cells such as dendritic cells that has an effect on the survival, proliferation, activation, and/or differentiation of other cell types, most notably other hematopoietic cells. Therefore, it may have clinical utility in treating a broad range of hematopoietic disorders and in increasing stem cell numbers. Furthermore, expression of this gene product in primary dendritic cells also strongly indicates a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1204 of SEQ ID NO:22, b is an integer of 15

to 1218, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

### **FEATURES OF PROTEIN ENCODED BY GENE NO: 13**

5           The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

          This gene is expressed primarily in fetal liver and spleen tissues, and to a lesser extent in breast tissue and Hodgkin's lymphoma.

10           Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune dysfunction; hematopoietic disorders; breast cancer; and Hodgkin's lymphoma. Similarly, polypeptides and antibodies directed to these  
15 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and/or breast, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., breast, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum,  
20 plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

          Preferred epitopes include those comprising a sequence shown in SEQ ID NO:  
25 56 as residues: Tyr-41 to Pro-46.

          The tissue distribution in fetal liver/spleen tissue, breast tissue, and Hodgkin's lymphoma, indicates that the protein products of this gene are useful for the diagnosis and/or treatment of a variety of hematopoietic disorders, including Hodgkin's lymphoma, as well as disorders of the breast, most notably breast cancer, as well as  
30 cancers of other tissues where expression has been observed. Expression of this gene product in hematopoietic tissues, particularly tissues involved in hematopoiesis such as fetal liver, suggest that it may play a role in the survival, proliferation, activation, and/or differentiation of hematopoietic lineages. Particularly, expression in Hodgkin's lymphoma indicates that it may be involved in proliferation and/or transformation,  
35 suggesting that it may also contribute to a variety of cancer processes. Expression in the breast indicates that it may be involved in normal breast function, in breast cancer, as a vital nutrient to infants during lactation, or may reflect expression within the lymph

nodes of the breast. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 698 of SEQ ID NO:23, b is an integer of 15 to 712, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a + 14.

#### 15 **FEATURES OF PROTEIN ENCODED BY GENE NO: 14**

The gene encoding the disclosed cDNA is thought to reside on chromosome 8. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 8.

This gene is expressed primarily in infant brain tissue, and to a lesser extent in T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system (CNS), expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 57 as residues: Ala-67 to Glu-72, Thr-91 to Ile-100.

The tissue distribution in infant brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of

neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in  
 5 feeding, sleep patterns, balance, and perception.

In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed  
 10 tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the  
 15 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1408 of SEQ ID NO:24, b is an integer of 15 to 1422, where both a and b correspond to the positions of nucleotide residues shown  
 20 in SEQ ID NO:24, and where b is greater than or equal to a + 14.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 15

The translation product of this gene shares sequence homology with phosphatidylethanolamine N-methyltransferase (isolated from rat) which is thought to  
 25 be important in catalyzing the synthesis of phosphatidylcholine from phosphatidylethanolamine in hepatocytes (See Genbank Accession No.: g310195 and J. Biol. Chem. 268 (22), 16655-16663 (1993)). Based on the sequence similarity between rat phosphatidylethanolamine N-methyltransferase and the translation product of this gene, the two proteins are expected to share certain biological activities.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: GGPRMKRSGNPGA EVTNSSVAGPDCCGGLGNIDFRQA  
 30 DFCVMTRLLGYVDPLDPSFVA AVITITTFNPLYWNVVARWEHKTRKLSRAFGSP  
 YLACYSLSXTILLN FLRSHCFTQA (SEQ ID NO:93); GGPRMKRSGNPGA EVT  
 NSSVAGPDCCGGLGNIDFRQADFCVMTRLLG YVDP (SEQ ID NO:94); and/or  
 35 LDPSFVA AVITITTFNPLYWNVVARWEHKTRKLSRAFGSPYLACYSLSXTILL  
 LN FLRSHCFTQA (SEQ ID NO:96). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is thought

to reside on chromosome 17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

This gene is expressed primarily in liver cells, and to a lesser extent in placental tissue.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, liver failure and liver metabolic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes  
10 for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and hepatic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., liver, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another  
15 tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 58 as residues: Pro-5 to Leu-10.

20 The tissue distribution in liver tissue, and the homology to phosphatidylethanolamine N-methyltransferase, indicates that the protein products of this gene are useful for the treatment and/or diagnosis of diseases of the liver, and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells).  
25 Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present  
30 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1024 of SEQ ID NO:25, b is an integer of 15  
35 to 1038, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 16**

The translation product of this gene shares sequence homology with heat shock protein 90, which is thought to be important in cellular proliferation. In specific

5 embodiments, polypeptides of the invention comprise the following amino acid sequences: PQRSELAASNRPCRVCISLLLCLEDRTMPKKAKPTGSGKEEGP APCKQMKLEAAGGPSALNFDSPSSLFESLISPIKTETFFKEFWEQKPLLIQRDD PALATYYGSLFKLTDLKSLCSRGMYYGRDVNVCRCVNGKKKVLNKGDKAHF LQLRKDFDQKRATIQFHQPQRFKDELWRIQEKLECYFGSLVGSNVYITPADLRA

10 CRPIMMMSRFSSCSWRERNTGASTTPLCPWHESTAWRPRKGSAGRCMSLC (SEQ ID NO:97); PQRSELAASNRPCRVCISLLLCLEDRTMPKKAKPTGSGKEE GP (SEQ ID NO:98); APCKQMKLEAAGGPSALNFDSPSSLFESLISPIKTETFFKE FWEQ (SEQ ID NO:99); KPLLIQRDDPALATYYGSLFKLTDLKSLCSRGMYYGR DVNVCRC (SEQ ID NO:100); VNGKKKVLNKGDKAHFLQLRKDFDQKRATIQF

15 HQPQRFKDELWRI (SEQ ID NO:101); QEKLECYFGSLVGSNVYITPADLRACRPI MMMSRFSSCSWRERN (SEQ ID NO:102); and/or TGASTTPLCPWHESTAWRPR KGSAGRCMSLC (SEQ ID NO:103). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is thought to reside on chromosome 3. Accordingly, polynucleotides related to this invention are

20 useful as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in placental tissue, and to a lesser extent in melanocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, systemic lupus erythematosus and other autoimmune diseases, acute leukemia, and developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above

30 tissues or cells, particularly of the immune and developing systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the

35 standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.



Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 59 as residues: His-13 to Leu-21, Glu-36 to Tyr-44, Thr-103 to Trp-109.

The tissue distribution in placental tissue, and the homology to heat shock protein 90, indicates that the protein products of this gene are useful for the treatment and/or diagnosis of systemic lupus erythematosus, since in SLE there is an overexpression of this protein, its surface localization and auto-antibodies to it have been observed. More generally, the tissue distribution in placental tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function.

Alternately, this gene product may be produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene product may be produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1892 of SEQ ID NO:26, b is an integer of 15 to 1906, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a + 14.

### **35 FEATURES OF PROTEIN ENCODED BY GENE NO: 17**

The translation product of this gene shares sequence homology with prostaglandin D synthetase, which is thought to be important in blood-tissue barriers.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: GGGIHLHNGALQLRVLQRVEHLHLLHHAVKHICTASLPVLHG FIAAQCRPGX (SEQ ID NO:104). Polynucleotides encoding these polypeptides are also encompassed by the invention.

5           This gene is expressed primarily in epididymus tissue.

          Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, multiple sclerosis, Meckel syndrome, polycystic kidney disease, and  
10       reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous, reproductive, and renal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell  
15       types (e.g., neural, renal, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20           The tissue distribution in epididymus tissue, and the homology to prostaglandin D synthetase, indicates that the protein products of this gene are useful for the treatment and/or diagnosis of diseases related to the blood-tissue, blood-cerebrospinal fluid, blood-retina, blood aqueous humor, and blood-testis barriers. More generally, the protein product of this gene, based upon its tissue distribution, is useful for the  
25       detection and or treatment of male reproductive disorders concerning dysfunction of the epididymus. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

          Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ  
30       ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of  
35       a-b, where a is any integer between 1 to 833 of SEQ ID NO:27, b is an integer of 15 to 847, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 18**

The translation product of this gene shares sequence homology with fructose transporter protein and other sugar transporter proteins. Based on the sequence  
 5 similarity to other sugar transporter proteins the translation product of this gene is expected to share certain biological activities with these proteins such as sugar transport activities. Such activities can be assayed by methods known to those of skill in the art.

When tested against fibroblast cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response gene 1) promoter  
 10 element. Thus, it is likely that this gene activates fibroblast cells, and to a lesser extent, in integumentary cells and tissues, through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jak-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

15 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: WDRWSDSALRRLRGSGDLAGELEELEEERAAACQGCRAR  
 RPWELFQHRALRRQVTSLVVLGSAMELCGNDVYAYASSVFRKAGVPEAKIQY  
 AIIGTGSCCELLTAVVSVSLEGALPPPALWGGTPRSSALNQFTLQKKKKKKKKK  
 KKKKKKK (SEQ ID NO:105); RRLRGSGDLAGELEELEEERAAACQGCRARRPW  
 20 ELFQH (SEQ ID NO:106); RQVTSLVVLGSAMELCGNDVYAYASSVF (SEQ ID  
 NO:107); and/or TGSCCELLTAVVSVSLEGALPPPALWGGTPRSSAL (SEQ ID  
 NO:108). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in endometrial stromal cells.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive and metabolic diseases and/or disorders, particularly diabetes. Similarly, polypeptides and antibodies directed to these polypeptides are  
 30 useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., reproductive, metabolic, and/or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,  
 35 synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the

expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 61 as residues: Phe-45 to Trp-50, Ala-52 to Pro-59, Ser-149 to Leu-154, Gly-219 to Cys-233.

The homology to sugar transporter proteins (particularly the GLUT5 protein) indicates that the protein products of this clone are useful for the treatment and/or diagnosis of sugar metabolism disorders such as diabetes. Further, polynucleotides and polypeptides of the present invention may be expressed in vivo by administration of the claimed polynucleotide and polypeptides (see Geneseq T66495-96) for treatment of diabetes, or expressed in a host cell to prepare a recombinant cell that secretes insulin in response to glucose and which can be administered to a patient to treat diabetes. Alternatively, the tissue distribution in endometrial stromal cells, combined with the detected EGR1 biological activity, suggests the protein is useful for the diagnosis, treatment, and/or prevention of reproductive and developmental diseases and/or disorders. The protein is useful in the treatment and/or detection of proliferative conditions. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 971 of SEQ ID NO:28, b is an integer of 15 to 985, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where b is greater than or equal to a + 14.

### **FEATURES OF PROTEIN ENCODED BY GENE NO: 19**

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: HELRLRKNTVKFSLYRHFKNLIFAVLASIVFMGWTTK TFRIAKCQSDW (SEQ ID NO:109). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in endometrial tumor tissue, and to a lesser extent in placental tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and reproductive diseases and/or disorders, particularly endometrial tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developmental, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 62 as residues: Pro-27 to Arg-33, Asp-41 to Ile-47, Thr-73 to Asp-85.

The tissue distribution in endometrial tumor tissue and placental tissue indicates that protein products of this gene are useful for the treatment, diagnosis, and/or prevention of endometrial tumors, as well as tumors of other tissues where expression has been observed. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, detection, and/or treatment of developmental disorders.

The relatively specific expression of this gene product in placental tissue and the endometrium indicates it may be a key player in the proliferation, maintenance, and/or differentiation of various cell types during development. It may also act as a morphogen to control cell and tissue type specification. Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. Expression within cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders.

Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present

invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 900 of SEQ ID NO:29, b is an integer of 15 to 914, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 20**

The translation product of this gene was shown to have homology to the conserved dolichyl-phosphate beta-glucosyltransferase from *Saccharomyces cerevisiae* and *S. pombe* (See Genebank Accession No. gil535141) which is important in protein trafficking, post-translational processing and modification of proteins, protein secretion, and stabilizing secreted proteins. Proteins involved in glycosylation events have uses which are well known in the art, and that supercede those mentioned above.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: WIPRAAGIRHEESIAQRSYFRTLL (SEQ ID NO:110); ADTN FTQETAMTMITPSSKLTLTGKNKSWSSSTAVAALELVDPGCRNSARGINCSAF LLPYSSHVWVPLSGVVPLCQRNQGHTVWVQIIYSRSSFTDVFISTR (SEQ ID NO:111); MTMITPSSKLTLTGKNKSWSSSTAVAA (SEQ ID NO:112); RGINCS AFLLPYSSHVWVPL (SEQ ID NO:113); and/or VVPLCQRNQGHTVWVQIIYSRS SF (SEQ ID NO:114). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in infant brain tissue, and to a lesser extent in ovarian cancer tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, developmental, metabolic, neural, and proliferative diseases and/or disorders, particularly multiple sclerosis, dementia, and ovarian cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., developmental, metabolic, proliferative, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 63 as residues: Gly-26 to Gln-32, Pro-42 to Ser-50.

The tissue distribution in infant brain tissue indicates that the protein products of this gene are useful for the treatment and/or diagnosis of defects or problems associated with developmental processes, particularly in the brain. The homology to dolichyl-phosphate beta-glucosyltransferase from *Saccharomyces cerevisiae* and *S. pombe* indicates that the protein plays a vital role in normal cellular and protein metabolism and is useful in treating proliferative disorders, in addition to, correcting metabolic deficiencies via gene therapy (i.e. protein may be required for proper conformation and stability of key secreted protein or enzyme and the stable insertion of the encoding gene into a stem cell may correct this deficit).

The expression within infant tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders (i.e. may inhibit key cell cycle regulators via inhibition of endogenous equivalent of present invention). Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA).

Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or

prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1169 of SEQ ID NO:30, b is an integer of 15 to 1183, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

15

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 21**

The translation product of this gene shares sequence homology with chicken ring zinc finger protein, which is thought to be important in the regulation of transcription. Zinc ring finger proteins have uses well known in the art, and which are described elsewhere herein. Briefly, the protein may be involved in inter-cellular communication and proliferation events, leading to migration or differentiation, and possibly apoptosis and cell death. The protein was subsequently cloned and sequenced by another group (See, for example, Lomax, M.I., Prim. Sens. Neuron (1998), which is hereby incorporated by reference, herein).

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: IRRLCDNFDIKVLNAQRAGYKAAIVHNVDSDDLISMGSN DIEVLKKIDIPSVFIGESSANSLKDEFTYEKGGHLILVPEFSLPLEYYLIPFLIIVGI CLILIVIFMITKFVQDRHRARRNRLRKDQLKKLPVHKFKKGDEYDVCAICLDEY EDGDKLRILPCSHAYHCKCVDPWLTCTKKTCPVCKQKVVP SQGDS DSDTDSS QEENEVTEHTPLLRPLASVSAQSFGALSESRS HQNMTESSDY EEDDNEDTDSSD AE (SEQ ID NO:122); NFDIKVLNAQRAGYKAAIVHNVDSDD (SEQ ID NO:115); VLKKIDIPSVFIGESSANSLKDEFTYEK (SEQ ID NO:116); PEFSLPLEYYLIPFL IIVGICLILIVIFMI (SEQ ID NO:117); TKFVQDRHRARRNRLRKDQLKKLPVHK FKKGDEY (SEQ ID NO:118); EDGDKLRILPCSHAYHCKCVDPWLTCT (SEQ ID NO:119); VVPSQGDSDSDTDSSQEENEVTEH (SEQ ID NO:120); and/or QSFGALS



ESRSHQNMTESSDYEEDDNEDT (SEQ ID NO:121). The gene encoding the disclosed cDNA is believed to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

This gene is expressed many adult and fetal tissues.

5       Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, many diseases such as developmental, immune, and neural diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides  
10       are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic system, central nervous system, immune system and others, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developmental, immune, neural, and  
15       cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20       Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 64 as residues: Asn-43 to Asp-49, Ser-71 to Ala-76, Pro-84 to Gly-91.

The tissue distribution in fetal tissues, combined with the homology to ring zinc proteins, indicates that the protein products of this gene are useful for treating and/or diagnosing diseases in the immune system, hematopoietic system and developmental  
25       disorders. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, and as nutritional supplements.

The protein product of this clone may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating  
30       activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating  
35       hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic

shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures.

Moreover, the expression within fetal tissue indicates this protein may play a  
5 role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell  
10 death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue  
15 differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available  
20 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more  
25 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2363 of SEQ ID NO:31, b is an integer of 15 to 2377, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

### 30 **FEATURES OF PROTEIN ENCODED BY GENE NO: 22**

The translation product of this gene shares sequence homology with kidney transporter, which is thought to be important in kidney function and dialysis (See Genebank Accession No: gil3831566 (AF057039)). This protein was subsequently cloned and sequenced by another group (See, for example, Reid, G., Kidney Blood  
35 Press. Res. 21 (2-4), 233-237 (1998), which is hereby incorporated by reference, herein).

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: AQCSIYLIQVIFGAVDLPAKLVGFLVINSLGRRPAQ (SEQ ID NO:123); GTVQHLPNPGDLWCCGPACQACGLPCHQLPGSPACPDGCTAAGRHLHPAQWGDTPGPVHCPNLSCCAGEGLSGCLLQLHLPVYWELYPTMIRQTGMGMG  
5 GSTMARVGSIVSPLVSMTAELYPSMPLFIYGAVPVAASAVTVLLPETLGQPLPDT  
VQDLESRKKGKQTRQQQEHQKYMVPLQASAEKNGL (SEQ ID NO:124); LPNPG  
DLWCCGPACQACGLPCHQ (SEQ ID NO:125), GCTAAGRHLHPAQWGDTPGPV  
HCPNL (SEQ ID NO:126); LHPVYWELYPTMIRQTGMGMG (SEQ ID NO:127);  
LVSMTAELYPSMPLFIY GAVPVA (SEQ ID NO:128); and/or PDTVQDLESRKKGKQ  
10 TRQQQEHQKYMVP (SEQ ID NO:129). Polynucleotides encoding these polypeptides  
are also encompassed by the invention. The gene encoding the disclosed cDNA is  
believed to reside on chromosome 1. Accordingly, polynucleotides related to this  
invention are useful as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in fetal brain, fetal kidney and adult kidney  
15 tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental diseases and/or disorders, particularly kidney and neural  
20 disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal and urologic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developmental,  
25 neural, renal, urogenital, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in fetal and adult kidney tissues, combined with the homology to kidney specific transporter, indicates that the protein products of this gene are useful for the treatment and/or diagnosis of renal and urologic disorders, as well as developmental disorders of the central nervous system. Moreover, the protein product of this gene could be used in the treatment and/or detection of kidney diseases including  
30 renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's  
35

Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome.

Alternatively, polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 781 of SEQ ID NO:32, b is an integer of 15 to 795, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 23**

The translation product of this gene shares sequence homology with the ubiquitin-specific protease, UBP2, (See Geneseq Accession No.R36730), which is thought to be important in metabolic processes, tissue repair, and wound healing.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: CLEAAMIEGEIESLHSENSGKSGQEHWFTELPPVLTFELS RFEFNQALGRPEKIHNKLEFPQVLYLDRYMHRNREITRIKREEIKRLKDYLTVL QQRLERYLSYGSGPKRFPLVDVLQYALEFASSKPVCTSPVDDIDASSPPSGSIPS QTL PSTTEQQGALSSELPSTSPSSVAAISSRSVIHKPFTQSRIPDLPMHPAPRHIT

EEELSVLESCLHRWRTEIENDTRDLQESISRIHRTIELMYSKSMIQVPYRLHAVL  
VHEGQANAGHYWAYIFDHRESRWMKYNDIAVTKSSWEELVRDSFGGYRNAS  
AYCLMYINDKAQFLIQEEFNKETGQPLVGIETLPPDLRDFVEEDNQRFKEKELEE  
WDAQLAQKALQEKLLASQKLRESESVTTAQAAGDPEYLEQPSRSDFSKHLKE  
5 ETIQITKASHEHEDKSPETVLQSAIKLEYARLVKLAQEDTPPETDYRLHHVVVY  
FIQNQAPKKIIEKTLLEQFGDRNLSFDERCHNIMKVAQAKLEMIKPEEVNLEEYE  
EWHQDYRKFRETTMYLIIGLENFQRESYIDSLFLICAYQNNKELLSKGLYRGH  
DEELISHYRRECLLKLNEQAALFESGEDREVNNGLIIMNEFTVPFLPLLVDDEM  
EEKDILAVEDMRNRWCSYLGQEMEPHLQEKLTDFLPKLLDCSMEIKSFHEPPK  
10 LPSYSTHELCECFARIMLSLSTPADGR (SEQ ID NO:130); MIEGEIESLHSENS  
GKSGQEHWFT (SEQ ID NO:131); FELSRFEFNQALGRPEKIHNKLEFP (SEQ ID  
NO:132); ITRIKREEIKRLKDYLTVLQQRLE (SEQ ID NO:133); PKRFPLVDVL  
QYALEFASSKPVCTSPV (SEQ ID NO:134); IPSQTLPTTEQQGALSSELPSTSPS  
(SEQ ID NO:135); SVIHKPFTQSRIPDLPMHPAPRH (SEQ ID NO:136); CLHRW  
15 RTEIENDTRDLQESISRI (SEQ ID NO:137); KSMIQVPYRLHAVLVHEGQANAG  
HYWAY (SEQ ID NO:138); RWMKYNDIAVTKSSWEELVRDSFGGYRNA (SEQ  
ID NO:139); INDKAQFLIQEEFNKETGQPLVGI (SEQ ID NO:140); MIQVPYRLHA  
VLVHEGQANAGHY (SEQ ID NO:141); DNQRFKEKELEEWDAQLAQKALQEKLL  
(SEQ ID NO:142); SETSVTTAQAAGDPEYLEQPSRS (SEQ ID NO:143); QITKA  
20 SHEHEDKSPETVLQSAIKLEYA (SEQ ID NO:144); LAQEDTPPETDYRLHHVVV  
YFIQNQAPK (SEQ ID NO:145); GDRNLSFDERCHNIMKVAQAKLEMIKPEE  
(SEQ ID NO:146); EEWQDYRKFRETTMYLIIGLENFQR (SEQ ID NO:147); CAY  
QNNKELLSKGLYRGHDEELISHYR (SEQ ID NO:148); CLLKLNEQAALFESG  
DREVNGLIIM (SEQ ID NO:149); VDEMEEKDILAVEDMRNRWCSYLGQEMEP  
25 HL (SEQ ID NO:150); and/or QEKLTDFLPKLLDCSMEIKSFHEPP (SEQ ID  
NO:151). Polynucleotides encoding these polypeptides are also encompassed by the  
invention. The gene encoding the disclosed cDNA is believed to reside on chromosome  
21. Accordingly, polynucleotides related to this invention are useful as a marker in  
linkage analysis for chromosome 21.

30 This gene is expressed primarily in fetal tissues and tumors thereof.

Therefore, polynucleotides and polypeptides of the invention are useful as  
reagents for differential identification of the tissue(s) or cell type(s) present in a  
biological sample and for diagnosis of diseases and conditions which include, but are  
not limited to, developmental diseases and/or disorders, particularly cancers. Similarly,  
35 polypeptides and antibodies directed to these polypeptides are useful in providing  
immunological probes for differential identification of the tissue(s) or cell type(s). For a  
number of disorders of the above tissues or cells, particularly of the immune system,

expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having  
5 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 66 as residues: Tyr-29 to Gln-46.

The tissue distribution in fetal tissues and tumors thereof, combined with the  
10 homology to a human ubiquitin-specific protease, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of cancers and developmental disorders. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, detection, and/or treatment of developmental disorders, and may be a key player in the proliferation, maintenance,  
15 and/or differentiation of various cell types during development. It may also act as a morphogen to control cell and tissue type specification. Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers.

Expression within fetal tissue and other cellular sources marked by proliferating  
20 cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and/or treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some  
25 cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may  
30 modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available  
35 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the

scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2642 of SEQ ID NO:33, b is an integer of 15 to 2656, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 24

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: QIATSVHHNINRKKRSVLRL (SEQ ID NO:152).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in fetal heart tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and cardiovascular diseases and/or disorders, particularly heart diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the heart, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developmental, cardiovascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 67 as residues: Ser-19 to Ser-25, Pro-27 to Gly-33, Pro-40 to Asn-47, Pro-65 to Gln-70.

The tissue distribution in fetal heart tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and/or treating heart diseases. The protein is useful in treating and/or detecting, but not limited to, the following: congenital birth defects, myocardial infarction, atherosclerosis, arteriosclerosis, endocarditis, cardiomyopathies, and myocarditis.

Moreover, the expression within fetal tissue indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues

rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain

5 neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of  
10 degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ  
15 ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of  
20 a-b, where a is any integer between 1 to 2552 of SEQ ID NO:34, b is an integer of 15 to 2566, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 25**

25 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: PLLRGLFIRXRAGHYECVFHEXVEGGACCEQC (SEQ ID NO:153); LVNNSFFLEFIYRPDSKNWQYQETIKKGDLLLNRVQKLSRVINM (SEQ ID NO:154); and/or IRELSRFIAAGRLHCKIDKVNEIVETNRYSHFSE (SEQ ID NO:155). Polynucleotides encoding these polypeptides are also encompassed by the  
30 invention.

This gene is expressed primarily in activated T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are  
35 not limited to, immune and hematopoietic disorders and/or diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a



number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 68 as residues: Gln-23 to Asn-28, Gly-38 to Ile-43.

The tissue distribution in activated T-cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of immune system disorders. Moreover, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the

scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1654 of SEQ ID NO:35, b is an integer of 15 to 1668, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 26

The translation product of this gene shares sequence homology with glutathione-S-transferase, which is thought to be important in inflammatory responses.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: GSQPPGPVPEXLIRIYSMRFCPYSHRTRLVLKAKDIRHE VVNINLRNKPEWYYTKHPFGHIPVLETSQCQLIYESVIACEYLDDAYPGRKLFP YDPYERARQKMLLELFCKVPHLTKECLVALRCGRECTNLKAALRQEFNSNLEEIL EYQNTTFFGGTCISMIDYLLWPWFERLDVYGILDCVSHTPACGSGYQP (SEQ ID NO:156); LASPPVPLHRCSA (SEQ ID NO:157); MRFCPYSHRTRLVLKAKDIRH EVVNINLR (SEQ ID NO:158); NKPEWYYTKHPFGHIPVLETSQCQ (SEQ ID NO:159); KLFPYDPYERARQKMLLELFCKVP (SEQ ID NO:160); VALRCGRECT NLKAALRQEFNSNLEE (SEQ ID NO:161); AAGCVWDTGLCEPHXSLRLWISAM KWDPTVCALLMDKSIFQGFLNLYFQNNPNAFDGLC (SEQ ID NO:163); and/or SMIDYLLWPWFERLDVYGILDCVS (SEQ ID NO:162). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in keratinocytes, melanocytes, and fetal skin tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, integumentary, inflammatory, and/or developmental diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., integumentary, inflammatory, developmental, metabolic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in integumentary cells and tissues, combined with the homology to glutathione-S-transferase, indicates that the protein products of this gene are useful for the diagnosis and treatment of inflammatory and skin diseases. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, urticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete's foot, and ringworm).

Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders such as arthritis, trauma, tendonitis, chondromalacia and inflammation, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:36 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 969 of SEQ ID NO:36, b is an integer of 15 to 983, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where b is greater than or equal to a + 14.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 27**

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: VYLFLTYRQAVVIALLVKVGVISEKHTWEWQTVEAVATG  
 5 LQDFIICIEMFLAAIAHHYTF SYKPYVQEAEEGSCFDSFLAMWDVSDIRDDISE  
 QVRHVGRTRVGRHPRKKLFPEDQDQNEHTSLLSSSSQDAISSMPPSPMGHY  
 QGFGHTVTPQTTPPTAK ISDEILSDTIGEEKEPS (SEQ ID NO:164); TNNKDSL  
 G WYLFTVLDSWIALKYPGIAIYVDTCRECYEAYVIYNFMGFLTNYLTNRYPNLVL  
 ILEAKDQQKHFPPLCCCPWAMGEVLLFRCKLSVLQYTVVRPFTTIVALICELG  
 10 IYDEGNFSFSNAWTYLVIINNMSQLFAMYCLLLFYKVLKEELSPIQPVGKFLCV  
 KLVVF (SEQ ID NO:165); and/or QNSQRTGLPITIFSRSFLLTGSDLCEN (SEQ ID  
 NO:166). Polynucleotides encoding these polypeptides are also encompassed by the  
 invention. The gene encoding the disclosed cDNA is believed to reside on chromosome  
 4. Accordingly, polynucleotides related to this invention are useful as a marker in  
 15 linkage analysis for chromosome 4.

This gene is expressed primarily in retinal tissue, and to a lesser extent in  
 keratinocytes, T-helper cells, endometrial tumor cells and infant brain tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as  
 reagents for differential identification of the tissue(s) or cell type(s) present in a  
 20 biological sample and for diagnosis of diseases and conditions which include, but are  
 not limited to, visual and immune diseases and/or disorders. Similarly, polypeptides  
 and antibodies directed to these polypeptides are useful in providing immunological  
 probes for differential identification of the tissue(s) or cell type(s). For a number of  
 disorders of the above tissues or cells, expression of this gene at significantly higher or  
 25 lower levels may be routinely detected in certain tissues or cell types (e.g., visual,  
 immune, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum,  
 plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken  
 from an individual having such a disorder, relative to the standard gene expression  
 level, i.e., the expression level in healthy tissue from an individual not having the  
 30 disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:  
 70 as residues: Thr-6 to Trp-13.

The tissue distribution is retinal tissue indicates that polynucleotides and  
 polypeptides corresponding to this gene are useful in the treatment and/or diagnosis of  
 35 visual disorders, which include, but are not limited to glaucoma, retinal/macular  
 degeneration, cataracts, conjunctivitis, and/or autoimmune disorders. Moreover, the  
 expression of this gene product in immune tissues indicates a role in regulating the

proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

5           Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as  
10 T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that  
15 influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor  
20 marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the  
25 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2337 of SEQ ID NO:37, b is an integer of 15 to 2351, where both a and b correspond to the positions of nucleotide residues shown  
30 in SEQ ID NO:37, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 28**

When tested against Jurkat and U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter  
35 element. Thus, it is likely that this gene activates promyelocytic and T-cells, and to a lesser extent, immune cell and tissues, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are

involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: QFFLCRDCS (SEQ ID NO:167); ERESCSHQAGVQWCNLSL RPPPPGFKQFSHLSPSS (SEQ ID NO:168); LRENALSSRLECSGAISAHCD LHLLGSSNSPTSASQVVRTTGAHHQAQPIFVFLVETGFHHVGQAHLKQLTSRY PPHLASQSAGITGMSYRTQPKLLWFYLYKQFKQYREVGSRK (SEQ ID NO:169); SSRLECSGAISAHCDLHLLGSSNSP (SEQ ID NO:170); GAHHQAQPIFVFLVET GFHHVGQAHLKQLTSRYPPHLASQ (SEQ ID NO:171); and/or ITGMSYRTQPKL LWFYLYKQFKQYR (SEQ ID NO:172). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in kidney tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renal and/or urogenital diseases and/or conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the kidney, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., renal, urogenital, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in kidney tissue, combined with the detected GAS biological activity, indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and/or treating kidney diseases. Moreover, the protein product of this gene could be used in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Alternatively, expression of this gene

product in the testis may implicate this gene product in normal testicular function. In addition, this gene product may be useful in the treatment of male infertility, and/or could be used as a male contraceptive. Moreover, conditions such as infertility and reduced sperm count can be assessed using the invention to determine whether the condition is associated with or caused by the occurrence of the gene or gene alteration. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1520 of SEQ ID NO:38, b is an integer of 15 to 1534, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 29**

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ENFPETREVRAFSPRENLELCTCKS (SEQ ID NO:173). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in K562 cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic diseases and/or conditions, particularly leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in K562 cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and/or treating leukemia. The protein product of this gene is useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex- vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1168 of SEQ ID NO:39, b is an integer of 15 to 1182, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where b is greater than or equal to a + 14.

25

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 30**

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ALYCSPSLQID (SEQ ID NO:174). Polynucleotides encoding these polypeptides are also encompassed by the invention.

30 This gene is expressed primarily in activated T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system,

35



expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in activated T-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and/or treating immune disorders. Moreover, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues.

Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of

a-b, where a is any integer between 1 to 1827 of SEQ ID NO:40, b is an integer of 15 to 1841, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where b is greater than or equal to a + 14.

## 5 FEATURES OF PROTEIN ENCODED BY GENE NO: 31

The translation product of this gene was shown to have homology to the human AF-6 gene product (See Genbank Accession No.gnllPIDId1033446 (AB011399)), which is thought to be important in the predisposition of acute myeloid leukemia.

In specific embodiments, polypeptides of the invention comprise the following  
 10 amino acid sequence: CHCSMLKSHGDTVQNVLTFLFVTLSYLVSDVSYLQIQKKLR (SEQ ID NO:175); and/or CYFHQKAQSNGPEKQEKEGVIQNFKRTLKKEK KEKKKK (SEQ ID NO:176). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 6. Accordingly, polynucleotides related to this invention are  
 15 useful as a marker in linkage analysis for chromosome 6.

This gene is expressed primarily in merkel cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are  
 20 not limited to, immune and hematopoietic disorders and/or diseases, particularly leukemias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may  
 25 be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, leukemic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not  
 30 having the disorder.

The tissue distribution in merkel cells, combined with the homology to the AF-6 gene, indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and/or treating immune disorders. The protein product of this gene is useful for the treatment and/or diagnosis of hematopoietic related disorders such  
 35 as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex- vivo culture, bone marrow transplantation, bone marrow

reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1183 of SEQ ID NO:41, b is an integer of 15 to 1197, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where b is greater than or equal to a + 14.

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
1	HCGMD59	209627 02/12/98	pCMVSPORT 2.0	11	790	1	780	438	438	44	1	30	31	74
2	HCNSD76	209627 02/12/98	pBluescript	12	554	1	554	134	134	45	1	29	30	77
3	HCNSD93	209627 02/12/98	pBluescript	13	1106	1	1106	139	139	46	1	21	22	46
4	HCWBE22	209627 02/12/98	ZAP Express	14	568	1	568	101	101	47	1	23	24	101
5	HFEAN33	209627 02/12/98	Uni-ZAP XR	15	3692	1	458	25	25	48	1	26	27	381
5	HFEAN33	209626 02/12/98	Uni-ZAP XR	42	602	1	602	25	25	75	1	26	27	177
6	HCWUM50	209627 02/12/98	ZAP Express	16	1428	208	1428	270	270	49	1	30	31	45
7	HDHIA94	209627 02/12/98	pCMVSPORT 2.0	17	1489	1	1489	154	154	50	1	30	31	168
7	HDHIA94	209627 02/12/98	pCMVSPORT 2.0	43	2492	1	2492	163	163	76	1	30	31	48
8	HDPAE76	209627 02/12/98	pCMVSPORT 3.0	18	1940	1	1940	25	25	51	1	31	32	49
9	HDP1054	209627 02/12/98	pCMVSPORT 3.0	19	1592	1	1592	71	71	52	1	29	30	40
10	HDPNC61	209627 02/12/98	pCMVSPORT 3.0	20	1410	1	1410	20	20	53	1	22	23	94

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
11	HDPND46	209627 02/12/98	pCMVSPORT 3.0	21	1727	1	1727	15	15	54	1	22	23	484
12	HDPSU13	209627 02/12/98	pCMVSPORT 3.0	22	1218	1	1218	14	14	55	1	25	26	114
13	HDTGC73	209627 02/12/98	pCMVSPORT 2.0	23	712	1	712	386	386	56	1	31	32	49
14	HE2PD49	209627 02/12/98	Uni-ZAP XR	24	1422	257	1404	337	337	57	1	18	19	171
15	HEEJ02	209627 02/12/98	Uni-ZAP XR	25	1038	148	1037	387	387	58	1	40	41	125
16	HELHD64	209627 02/12/98	Uni-ZAP XR	26	1906	538	1906	549	549	59	1	14	15	310
17	HEPAD91	209627 02/12/98	Uni-ZAP XR	27	847	1	847	161	161	60	1	20	21	163
18	HEQBH65	209627 02/12/98	pCMVSPORT 3.0	28	985	1	985	18	18	61	1	24	25	239
19	HETCO02	209627 02/12/98	Uni-ZAP XR	29	914	1	914	150	150	62	1	29	30	129
20	HFAUO78	209627 02/12/98	Uni-ZAP XR	30	1183	212	1183	360	360	63	1	21	22	60
21	HFKEE48	209627 02/12/98	Uni-ZAP XR	31	2377	137	1596	166	166	64	1	34	35	97
22	HFKFG02	209627 02/12/98	Uni-ZAP XR	32	795	1	795	110	110	65	1	18	19	53
23	HFPCN45	209627 02/12/98	Uni-ZAP XR	33	2656	291	2656	362	362	66	1	28	29	63

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
24	HHFFJ48	209627 02/12/98	Uni-ZAP XR	34	2566	1	2566	65	65	67	1	21	22	106
25	HILCF66	209627 02/12/98	pBluescript SK-	35	1668	740	1668	331	331	68	1	21	22	44
26	HKABN45	209627 02/12/98	pCMVSPORT 2.0	36	983	1	983	347	347	69	1	19	20	42
27	HKAEV06	209627 02/12/98	pCMVSPORT 2.0	37	2351	1	2351	197	197	70	1	29	30	57
28	HKDBK22	209627 02/12/98	pCMVSPORT 1	38	1534	1	1534	130	130	71	1	44	45	44
29	HKFBB67	209627 02/12/98	ZAP Express	39	1182	1	1182	231	231	72	1	33	34	70
30	HKGZ06	209627 02/12/98	pSport1	40	1841	1	1841	67	67	73	1	28	29	43
31	HKGCK61	209627 02/12/98	pSport1	41	1197	1	1197	182	182	74	1	20	21	42

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below).



It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

5       The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, *Gene* 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources  
10       using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

### Signal Sequences

Methods for predicting whether a protein has a signal sequence, as well as the  
15       cleavage point for that sequence, are available. For instance, the method of McGeoch, *Virus Res.* 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, *Nucleic Acids Res.* 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1  
20       indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, *supra*.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide  
25       was analyzed by a computer program called SignalP (Henrik Nielsen et al., *Protein Engineering* 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results  
30       shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., +  
35       or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely

uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

## 10 **Polynucleotide and Polypeptide Variants**

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

15 By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragement specified as described herein.

25 As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are:

30 Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization

Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query

amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions,

5 interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be  
10 determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and  
15 subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window  
20 Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity.  
25 For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of  
30 the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are  
35 considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired  
5 residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence.  
10 This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query  
15 sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or  
20 activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in  
25 the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level.  
30 Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be  
35 deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., *J. Biol. Chem.* 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after

deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of the present invention having an amino acid

sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a polypeptide to have an amino acid sequence which comprises the amino acid sequence of the present invention, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of the present invention or fragments thereof (e.g., the mature form and/or other fragments described herein), is 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

#### **Polynucleotide and Polypeptide Fragments**

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger



or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid  
5 sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40,  
10 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

15 Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the  
20 mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

25 Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-  
30 forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active  
35 fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

### Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')<sub>2</sub> fragments) which are capable of specifically binding to protein. Fab and F(ab')<sub>2</sub> fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

### **Fusion Proteins**

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively,

deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., *J. Molecular Recognition* 8:52-58 (1995); K. Johanson et al., *J. Biol. Chem.* 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., *Proc. Natl. Acad. Sci. USA* 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., *Cell* 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

## **Vectors, Host Cells, and Protein Production**

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli* lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a

translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera Sf9* cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., *Basic Methods In Molecular Biology* (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production

procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

### **Uses of the Polynucleotides**

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be

selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

5 Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-  
10 sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides  
15 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for  
20 marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the  
25 physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per  
30 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or  
35 translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the

mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

5        Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

10        In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids  
15 Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991) ) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off  
20 of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

      Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective  
25 gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

      The polynucleotides are also useful for identifying individuals from minute  
30 biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags"  
35 which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.



The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

### **Uses of the Polypeptides**

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in

tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine ( $^{125}\text{I}$ ,  $^{121}\text{I}$ ), carbon ( $^{14}\text{C}$ ), sulfur ( $^{35}\text{S}$ ), tritium ( $^3\text{H}$ ), indium ( $^{112}\text{In}$ ), and technetium ( $^{99\text{m}}\text{Tc}$ ), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example,  $^{131}\text{I}$ ,  $^{112}\text{In}$ ,  $^{99\text{m}}\text{Tc}$ ), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of  $^{99\text{m}}\text{Tc}$ . The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

### **Biological Activities**

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

### **Immune Activity**

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g.,

by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

5 A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to:

10 blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

15 Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet

20 disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

25 A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the

30 present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia,

35 antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus,

Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune inflammatory eye disease.

5 Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

10 A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits  
15 an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory  
20 response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel  
25 disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

### **Hyperproliferative Disorders**

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative  
30 disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing  
35 antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by

initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

- 5 Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.
- 10 Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenström's Macroglobulinemia, Gaucher's Disease, histiocytosis, and
- 15 any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

### **Infectious Disease**

- 20 A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the
- 25 infectious agent, without necessarily eliciting an immune response.

- Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus,
- 30 Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g.,
- 35 Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye

infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually  
 5 transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention  
 10 include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella,  
 15 Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Pasteurellaceae Infections (e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases  
 20 or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning,  
 25 Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect  
 30 any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis,  
 35 Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery,

giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

5            Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be  
10        used as an antigen in a vaccine to raise an immune response against infectious disease.

### **Regeneration**

          A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See,  
15        Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

20            Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

25            Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel  
30        syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

          Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate  
35        nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and



stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-  
5 Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

### **Chemotaxis**

10 A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

15 A polynucleotide or polypeptide of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat wounds and other trauma to  
20 tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used  
25 as an inhibitor of chemotaxis.

### **Binding Activity**

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding  
30 of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the  
35 polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural

receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate  
5 cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the  
10 polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

15 Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a  
20 standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

25 All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

30 Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with  
35 a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

### **Other Activities**

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as  
5 discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be  
10 used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, circadian rhythms, depression (including depressive disorders), tendency for violence, tolerance  
15 for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat  
20 content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

### **Other Preferred Embodiments**

Other preferred embodiments of the claimed invention include an isolated  
25 nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of  
30 positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of  
35 positions beginning with the nucleotide at about the position of the 5' Nucleotide of the

Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the  
5 range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide  
10 sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

15 A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in  
20 Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under  
25 stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which  
30 comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide  
35 sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA

Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any

integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

5 The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

10 Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence  
15 of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

20 The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

25 Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human  
30 cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the  
35 amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions

beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at  
5 least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

10 Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone  
15 identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and  
20 contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA  
25 clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA  
30 clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and  
35 contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded



by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide  
5 molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition  
10 associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a  
15 sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

20 In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least  
25 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA  
30 clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide  
35 comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone

identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

## **Examples**

### **Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample**

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the

related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

	<u>Vector Used to Construct Library</u>	<u>Corresponding Deposited Plasmid</u>
5	Lambda Zap	pBluescript (pBS)
	Uni-Zap XR	pBluescript (pBS)
	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
10	pCMVSPORT 2.0	pCMVSPORT 2.0
	pCMVSPORT 3.0	pCMVSPORT 3.0
	pCR <sup>®</sup> 2.1	pCR <sup>®</sup> 2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSPORT 2.0 and pCMVSPORT 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR<sup>®</sup>2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the

phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with  $^{32}\text{P}$ - $\gamma$ -ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25  $\mu\text{l}$  of reaction mixture with 0.5  $\mu\text{g}$  of the above cDNA template. A convenient reaction mixture is 1.5-5 mM  $\text{MgCl}_2$ , 0.01% (w/v) gelatin, 20  $\mu\text{M}$  each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation

at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence  
5 by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known  
10 in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set  
15 containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with  
20 phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction  
25 leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to  
30 the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

**Example 2: Isolation of Genomic Clones Corresponding to a**  
35 **P lynucleotide**

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

5     **Example 3: Tissue Distribution of Polypeptide**

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P<sup>32</sup> using the rediprime™ DNA labeling system  
10     (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or  
15     human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

20

**Example 4: Chromosomal Mapping of the Polynucleotides**

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of  
25     conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is  
30     determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

**Example 5: Bacterial Expression of a Polypeptide**

A polynucleotide encoding a polypeptide of the present invention is amplified  
35     using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA

sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance ( $Amp^r$ ), a bacterial origin of replication ( $ori$ ), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the *E. coli* strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance ( $Kan^r$ ). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.<sup>600</sup>) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the

protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After  
5 renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4° C or frozen at -80° C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively  
10 linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an *E. coli* origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (*lacIq*). The origin  
15 of replication (*oriC*) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with *NdeI* and *XbaI*, *BamHI*, *XhoI*, or *Asp718*, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA  
20 insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for *NdeI* (5' primer) and *XbaI*, *BamHI*, *XhoI*, or *Asp718* (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to  
25 express protein in a bacterial system.

#### **Example 6: Purification of a Polypeptide from an Inclusion Body**

The following alternative method can be used to purify a polypeptide expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified,  
30 all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount  
35 of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50



mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant  $A_{280}$  monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

### **Example 7: Cloning and Expression of a Polypeptide in a Baculovirus**

#### **Expression System**

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak *Drosophila* promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

5 The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

10 The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

15 Five µg of a plasmid containing the polynucleotide is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid are mixed in a sterile well of a  
20 microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then  
25 incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life  
30 Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a  
35 micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in

35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5  $\mu$ Ci of  $^{35}$ S-methionine and 5  $\mu$ Ci  $^{35}$ S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

#### **Example 8: Expression of a Polypeptide in Mammalian Cells**

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIV and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

5           The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and  
10       Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a  
15       chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No.209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et  
20       al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse  
25       DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol  
30       outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially  
35       available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five  $\mu$ g of the expression plasmid pC6 is cotransfected with 0.5  $\mu$ g of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1  $\mu$ M, 2  $\mu$ M, 5  $\mu$ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200  $\mu$ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

#### **Example 9: Protein Fusions**

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., *Nature* 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the half-life time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

5 For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that  
10 the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a  
15 heterologous signal sequence. (See, e.g., WO 96/34891.)

#### Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAACTCACACATGCCCACCGTGCC  
CAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCAAACC  
20 CAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGTGGT  
GGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACG  
GCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAAC  
AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTG  
AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAACCCCC  
25 ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGT  
GTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCT  
GACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGTGGAGTGGGA  
GAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTGG  
ACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCA  
30 GGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC  
ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGC  
GACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

#### **Example 10: Production of an Antibody from a Polypeptide**

35 The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera

containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

5 In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., *Nature* 256:495 (1975); Köhler et al., *Eur. J. Immunol.* 6:511 (1976); Köhler et al., *Eur. J. Immunol.* 6:292 (1976); Hammerling et al., in: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures  
10 involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at  
15 about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line  
20 (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (*Gastroenterology* 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

25 Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a  
30 mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific  
35 antibodies.



It will be appreciated that Fab and F(ab')<sub>2</sub> and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')<sub>2</sub> fragments). Alternatively,  
5 secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies  
10 described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268  
15 (1985).)

#### **Example 11: Production Of Secreted Protein For High-Throughput Screening Assays**

The following protocol produces a supernatant containing a polypeptide to be  
20 tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates)  
25 and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

30 Plate 293T cells (do not carry cells past P+20) at  $2 \times 10^5$  cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine  
35 (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in

Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl<sub>2</sub> (anhyd); 0.00130 mg/L CuSO<sub>4</sub>·5H<sub>2</sub>O; 0.050 mg/L of Fe(NO<sub>3</sub>)<sub>3</sub>·9H<sub>2</sub>O; 0.417 mg/L of FeSO<sub>4</sub>·7H<sub>2</sub>O; 311.80 mg/L of KCl; 28.64 mg/L of MgCl<sub>2</sub>; 48.84 mg/L of MgSO<sub>4</sub>; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO<sub>3</sub>; 62.50 mg/L of NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O; 71.02 mg/L of Na<sub>2</sub>HPO<sub>4</sub>; .4320 mg/L of ZnSO<sub>4</sub>·7H<sub>2</sub>O; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitic Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H<sub>2</sub>O; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H<sub>2</sub>O; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H<sub>2</sub>O; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalanine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2H<sub>2</sub>O; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B<sub>12</sub>; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine;

0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

#### **Example 12: Construction of GAS Reporter Construct**

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are  
5 generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-  
10 12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

15 Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the  
20 proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

111

<u>Ligand</u>	<u>tyk2</u>	<u>JAKs</u> <u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>	<u>STATS</u>	<u>GAS(elements) or ISRE</u>
<u>IFN family</u>						
IFN- $\alpha$ /B	+	+	-	-	1,2,3	ISRE
IFN-g		+	+	-	1	GAS (IRF1>Lys6>IFP)
IL-10	+	?	?	-	1,3	
<u>gp130 family</u>						
IL-6 (Pleiotrophic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
IL-11 (Pleiotrophic)	?	+	?	?	1,3	
OnM (Pleiotrophic)	?	+	+	?	1,3	
LIF (Pleiotrophic)	?	+	+	?	1,3	
CNTF (Pleiotrophic)	+/-	+	+	?	1,3	
G-CSF (Pleiotrophic)	?	+	?	?	1,3	
IL-12 (Pleiotrophic)	+	-	+	+	1,3	
<u>g-C family</u>						
IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP >> Ly6)(IgH)
IL-7 (lymphocytes)	-	+	-	+	5	GAS
IL-9 (lymphocytes)	-	+	-	+	5	GAS
IL-13 (lymphocyte)	-	+	?	?	6	GAS
IL-15	?	+	?	+	5	GAS
<u>gp140 family</u>						
IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)
IL-5 (myeloid)	-	-	+	-	5	GAS
GM-CSF (myeloid)	-	-	+	-	5	GAS
<u>Growth hormone family</u>						
GH	?	-	+	-	5	
PRL	?	+/-	+	-	1,3,5	
EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)
<u>Receptor Tyrosine Kinases</u>						
EGF	?	+	+	-	1,3	GAS (IRF1)
PDGF	?	+	+	-	1,3	
CSF-1	?	+	+	-	1,3	GAS (not IRF1)

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:  
 5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCG  
 10 AAATGATTTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTGGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:  
 5':CTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAATG  
 20 ATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCC  
 CTAATCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTTCTCCGC  
 CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGC  
 CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTT  
 TGCAAAAAGCTT:3' (SEQ ID NO:5)

25 With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase,  
 30 alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a  
 35 neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

**Example 13: High-Throughput Screening Assay for T-cell Activity.**

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies)

with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells ( $10^7$  per transfection), and resuspend in OPTI-MEM to a final  
5 concentration of  $10^7$  cells/ml. Then add 1ml of  $1 \times 10^7$  cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

10 On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

15 Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12  
20 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples  
25 from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

30 As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

The above protocol may be used in the generation of both transient, as well as, stable transfected cells, which would be apparent to those of skill in the art.



**Example 14: High-Throughput Screening Assay Identifying Myeloid Activity**

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells.

5 Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

10 To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest  $2 \times 10^7$  U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

15 Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ , 1 mM  $\text{MgCl}_2$ , and 675 uM  $\text{CaCl}_2$ . Incubate at 37°C for 45 min.

20 Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

25 These cells are tested by harvesting  $1 \times 10^8$  cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of  $5 \times 10^5$  cells/ml. Plate 200 ul cells per well in the 96-well plate (or  $1 \times 10^5$  cells/well).

30 Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

**Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.**

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat pheochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

5        The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as  $5 \times 10^5$  cells/ml.

10        Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to  $1 \times 10^5$  cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

15

**Example 16: High-Throughput Screening Assay for T-cell Activity**

NF- $\kappa$ B (Nuclear Factor  $\kappa$ B) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by  
20        expression of certain viral gene products. As a transcription factor, NF- $\kappa$ B regulates the expression of genes involved in immune cell activation, control of apoptosis (NF- $\kappa$ B appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF-  $\kappa$ B is retained in the cytoplasm with I- $\kappa$ B  
25        (Inhibitor  $\kappa$ B). However, upon stimulation, I-  $\kappa$ B is phosphorylated and degraded, causing NF-  $\kappa$ B to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF-  $\kappa$ B include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

30        Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- $\kappa$ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- $\kappa$ B would be useful in treating

diseases. For example, inhibitors of NF- $\kappa$ B could be used to treat those diseases related to the acute or chronic activation of NF- $\kappa$ B, such as rheumatoid arthritis.

To construct a vector containing the NF- $\kappa$ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- $\kappa$ B binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:  
 5':GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGAC  
 TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:  
 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)

Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGACTTTCC  
 ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCCA  
 TCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACT  
 AATTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTATTC  
 CAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT:  
 3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF- $\kappa$ B/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- $\kappa$ B/SV40/SEAP cassette is removed from the above NF- $\kappa$ B/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF- $\kappa$ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF- $\kappa$ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

#### **Example 17: Assay for SEAP Activity**

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15  $\mu$ l of 2.5x dilution buffer into Optiplates containing 35  $\mu$ l of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50  $\mu$ l Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50  $\mu$ l Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

#### **Reaction Buffer Formulation:**

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6

23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

**Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability**

5 Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a  
10 fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-4 (Molecular Probes, Inc.;  
15 catalog no. F-14202), used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO<sub>2</sub> incubator for 20 hours.

The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

5 A stock solution of 1 mg/ml fluo-4 is made in 10% pluronic acid DMSO. To load the cells with fluo-4, 50 ul of 12 ug/ml fluo-4 is added to each well. The plate is incubated at 37°C in a CO<sub>2</sub> incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10<sup>6</sup> cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-4 solution in 10% pluronic acid DMSO is added to each ml of cell suspension.

10 The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10<sup>6</sup> cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-4. The supernatant is added to the well, and a change in fluorescence is detected.

15

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca<sup>++</sup> concentration.

20

#### **Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity**

25 The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also

30 membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-

35 receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members

of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford, MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford, MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na<sub>3</sub>VO<sub>4</sub>, 2 mM Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.



Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and  
5 PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg<sub>2+</sub> (5mM ATP/50mM MgCl<sub>2</sub>), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride,  
10 pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl<sub>2</sub>, 5 mM MnCl<sub>2</sub>, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initiate the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mM  
15 EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-  
20 POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the  
25 absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

#### **Example 20: High-Throughput Screening Assay Identifying** 30 **Phosphorylation Activity**

As a potential alternative and/or complement to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine  
35 phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase,

Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

#### **Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide**

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

#### **Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample**

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with

specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

- 5       The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

- 10       Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

- 15       Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

## 20    **Example 23: Formulating a Polypeptide**

- The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of  
25   administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

- As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject  
30   to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous  
35   bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D-(-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's

solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

**Example 24: Method of Treating Decreased Levels of the Polypeptide**

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

**Example 25: Method of Treating Increased Levels of the Polypeptide**

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

**Example 26: Method of Treatment Using Gene Therapy**

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

5 pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

10 The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to  
15 transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is  
20 then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media,  
25 containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is  
30 required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.



**Example 27: Method of Treatment Using Gene Therapy - In Vivo**

Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide. The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622, 5705151, 5580859; Tabata H. et al. (1997) Cardiovasc. Res. 35(3):470-479, Chao J et al. (1997) Pharmacol. Res. 35(6):517-522, Wolff J.A. (1997) Neuromuscul. Disord. 7(5):314-318, Schwartz B. et al. (1996) Gene Ther. 3(5):405-411, Tsurumi Y. et al. (1996) Circulation 94(12):3281-3290 (incorporated herein by reference).

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Felgner P.L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain

sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies  
5 have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder,  
10 stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is  
15 similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and  
20 expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body  
25 weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being

treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15  $\mu$ m cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be used to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

**Example 28: Transgenic Animals.**

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, *e.g.*, baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (*i.e.*, polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., Appl. Microbiol. Biotechnol. 40:691-698 (1994); Carver et al., Biotechnology (NY) 11:1263-1270 (1993); Wright et al., Biotechnology (NY) 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus mediated gene transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56:313-321 (1989)); electroporation of cells or embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, *e.g.*, Ulmer et al., Science 259:1745 (1993); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (Lavitano et al., Cell 57:717-723 (1989); etc. For a review of such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115:171-229 (1989), which is incorporated by reference herein in its entirety.

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to

quiescence (Campell et al., Nature 380:64-66 (1996); Wilmut et al., Nature 385:810-813 (1997)).

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their  
5 cells, *i.e.*, mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, *e.g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89:6232-6236 (1992)). The regulatory  
10 sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the  
15 endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., Science  
20 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may  
25 be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and reverse

transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

#### **Example 29: Knock-Out Animals.**

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (*E.g.*, see Smithies et al., *Nature* 317:230-234 (1985); Thomas & Capecchi, *Cell* 51:503-512 (1987); Thompson et al., *Cell* 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or

regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention *in vivo*. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest.

5 Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (*e.g.*, see Thomas & Capecchi 1987 and Thompson 1989, *supra*). However this approach can be routinely adapted

10 for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically

15 engineered not to express the polypeptides of the invention (*e.g.*, knockouts) are administered to a patient *in vivo*. Such cells may be obtained from the patient (*i.e.*, animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (*e.g.*, lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered *in vitro* using

20 recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, *e.g.*, by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited

25 to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc. The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the

invention can be introduced into the patient systemically, e.g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft; 5 genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Patent No. 5,399,349; and Mulligan & Wilson, U.S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).

When the cells to be administered are non-autologous or non-MHC compatible 10 cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

15 Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

20

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

25 The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both 30 incorporated herein by reference in their entireties.



## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A.</b> The indications made below relate to the microorganism referred to in the description on page <u>54</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> <span style="float: right;">Further deposits are identified on an additional sheet <input type="checkbox"/></span>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution <i>(including postal code and country)</i> <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>February 12, 1998</u>	Accession Number <u>209627</u>
<b>C. ADDITIONAL INDICATIONS</b> <i>(leave blank if not applicable)</i> <span style="float: right;">This information is continued on an additional sheet <input type="checkbox"/></span>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> <i>(if the indications are not for all designated States)</i>	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> <i>(leave blank if not applicable)</i>	
The indications listed below will be submitted to the International Bureau later <i>(specify the general nature of the indications e.g., "Accession Number of Deposit")</i>	

<p style="text-align: center;"><b>For receiving Office use only</b></p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <hr/> <p>Authorized officer: <u>Lydell Meadows</u>  <u>Paralegal Specialist</u>  <u>APD-PCT Operations</u></p>	<p style="text-align: center;"><b>For International Bureau use only</b></p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <hr/> <p>Authorized officer</p>
--	---

*What Is Claimed Is:*

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
- (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
  - (f) a polynucleotide which is a variant of SEQ ID NO:X;
  - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
  - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
  - (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.
2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.
- 5
5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
- 10
6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
- 15
7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.
8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.
- 20
9. A recombinant host cell produced by the method of claim 8.
10. The recombinant host cell of claim 9 comprising vector sequences.
11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:
- 25
- (a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- (b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;
- 30
- (c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- 35
- (e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- (f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

- (g) a variant of SEQ ID NO: Y;
- (h) an allelic variant of SEQ ID NO: Y; or
- (i) a species homologue of the SEQ ID NO: Y.

5 12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

10 13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.

14. A recombinant host cell that expresses the isolated polypeptide of claim 11.

15 15. A method of making an isolated polypeptide comprising:  
(a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and  
(b) recovering said polypeptide.

20 16. The polypeptide produced by claim 15.

17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.

25 18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and  
(b) diagnosing a pathological condition or a susceptibility to a pathological  
30 condition based on the presence or absence of said mutation.

19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or amount of expression of the polypeptide of  
35 claim 11 in a biological sample; and  
(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

- 5 (a) contacting the polypeptide of claim 11 with a binding partner; and  
(b) determining whether the binding partner effects an activity of the polypeptide.

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

10 22. A method of identifying an activity in a biological assay, wherein the method comprises:

- (a) expressing SEQ ID NO:X in a cell;  
(b) isolating the supernatant;  
(c) detecting an activity in a biological assay; and  
15 (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 20.

<110> Human Genome Sciences, Inc. et al.

<120> 31 Human secreted proteins

<130> PZ026PCT

<140> Unassigned

<141> 1999-03-11

<150> 60/077,714

<151> 1998-03-12

<150> 60/077,687

<151> 1998-03-12

<150> 60/077,686

<151> 1998-03-12

<150> 60/077,696

<151> 1998-03-12

<160> 176

<170> PatentIn Ver. 2.0

<210> 1

<211> 733

<212> DNA

<213> Homo sapiens

<400> 1

gggatccgga	gccccaaatct	tctgacaaaa	ctcacacatg	cccaccgtgc	ccagcacctg	60
aattcgagg	tgcaccgtca	gtcttctct	tcccccaaa	acccaaggac	accctcatga	120
tctcccggac	tcttgaggtc	acatgcgtgg	tggtggacgt	aagccacgaa	gaccctgagg	180
tcaagttcaa	ctggtacgtg	gacggcgtgg	aggtgcataa	tgccaagaca	aagccgcggg	240
aggagcagta	caacagcacg	taccgtgtgg	tcagcgtcct	caccgtcctg	caccaggact	300
ggctgaatgg	caaggagtac	aagtgcgaag	tctccaacaa	agccctccca	acccccatcg	360
agaaaaccat	ctccaaagcc	aaagggcagc	cccgagaacc	acaggtgtac	accctgcccc	420
catcccggga	tgagctgacc	aagaaccagg	tcagcctgac	ctgcctggtc	aaaggcttct	480
atccaagcga	catcgccgtg	gagtgggaga	gcaatgggca	gccggagAAC	aactacaaga	540
ccacgcctcc	cgtgctggac	tccgacggct	ccttcttct	ctacagcaag	ctcaccgtgg	600
acaagagcag	gtggcagcag	gggaacgtct	tctcatgtc	cgtgatgcat	gaggctctgc	660
acaaccacta	cacgcagaag	agcctctccc	tgtctccggg	taaagagtg	cgacggccgc	720
gactctagag	gat					733

<210> 2

<211> 5

<212> PRT

<213> Homo sapiens

<220>

<221> Site

<222> (3)

<223> Xaa equals any of the twenty naturally occurring L-amino acids

<400> 2

Trp Ser Xaa Trp Ser

1

5

<210> 3

<211> 86

<212> DNA

<213> Homo sapiens

<400> 3

gcgctcgcag atttccccga aatctagatt tccccgaaat gatttccccg aaatgatttc 60  
cccgaaatat ctgccatctc aattag 86

<210> 4

<211> 27

<212> DNA

<213> Homo sapiens

<400> 4

gcggcgaagct ttttgcaaag cctaggc 27

<210> 5

<211> 271

<212> DNA

<213> Homo sapiens

<400> 5

ctcgagattt ccccgaaatc tagatttccc cgaaatgatt tccccgaaat gatttccccg 60  
aaatatctgc catctcaatt agtcagcaac catagtcccg ccctaactc cgcccatccc 120  
gcccctaact ccgccagtt ccgccattc tccgcccacat ggctgactaa ttttttttat 180  
ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 240  
ttttggaggc ctaggctttt gcaaaaagct t 271

<210> 6

<211> 32

<212> DNA

<213> Homo sapiens

<400> 6

gcgctcgcagg gatgacagcg atagaacccc gg 32

<210> 7

<211> 31

<212> DNA

<213> Homo sapiens

<400> 7

gcgaagcttc gcgactcccc ggatccgcct c 31

<210> 8  
 <211> 12  
 <212> DNA  
 <213> Homo sapiens

<400> 8  
 ggggactttc cc 12

<210> 9  
 <211> 73  
 <212> DNA  
 <213> Homo sapiens

<400> 9  
 gcggcctcga ggggactttc ccggggactt tccggggact ttccgggact ttccatcctg 60  
 ccattctcaat tag 73

<210> 10  
 <211> 256  
 <212> DNA  
 <213> Homo sapiens

<400> 10  
 ctcgagggga ctttcccgga gactttccgg ggactttccg ggactttcca tctgccatct 60  
 caattagtag gcaaccatag tcccgccctt aactccgccc atcccgcccc taactccgcc 120  
 cagttccgcc cattctccgc cccatggctg actaattttt tttatttatg cagaggccga 180  
 ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggcttttttg gaggcctagg 240  
 cttttgcaaa aagctt 256

<210> 11  
 <211> 790  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (37)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (55)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (76)  
 <223> n equals a,t,g, or c

<220>



<221> SITE  
 <222> (112)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (120)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (137)  
 <223> n equals a,t,g, or c

<400> 11  
 tcaactgggt gaaaaggaaa acccaccctt ggcgccnaat acgcaaaccg ccttntcccc 60  
 ggcgcggttg ccgatncatt aatgcagctg gcacgacagt tttcccgact gnaaagcggn 120  
 cagtgcgcgc aacgcantta aatgtgagtt agctcactca ttagcacccc aggctttaca 180  
 ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc ggataacaat ttcacacagg 240  
 aaacagctat gaccatgatt acgccaagct ctaatacgac tactatagg gaaagctggt 300  
 acgcctgcag gtaccgggtcc ggaattcccc ggtcgaccca cgcgccgggt tgaatgcact 360  
 gagtcccttg gtgtagtagc aataaggaaa aatgaaatta ctttcctgtg cacacagtcc 420  
 agcctaattg gtatgtgatg ttgcacttag cagccatgtg gtgggcatgt gtgactactc 480  
 tggttttcac ttttagtttct aaacttttta tccctctcaa gtccagcatg gatggggaaa 540  
 tgtctctgga tccccacagc tgtgtacttg tttgcatttg tttccctttg agatttgtgt 600  
 ttgtgtcctg ctttgagctg taccttgtcc agtccattgt gaaattatcc cagcagctgt 660  
 aatgtacagt tccttctgaa gcaagcaaca tcagcagcag cagcagcagc agcacaattc 720  
 tgtgttttat aaagacaaca gtggcttcta wwaaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 780  
 aaaaaaaaaa 790

<210> 12  
 <211> 554  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (552)  
 <223> n equals a,t,g, or c

<400> 12  
 ttcggcacga ggtctttacc tccaaactaa cttctttcct gaacagtaga atagtttttc 60  
 atactatcat catattggatg gagctcttta aactgacctc agagatcaga ttcataacct 120  
 tttgtccaga gcaatggatg cctttgctgg ttccccgttc tcattgatgg tccctaaatg 180  
 tgtacttata ctgttctgtc tagtctacag cttacagtgc attcagcctt attcaagctt 240  
 attgaattca gcctcggtgc cttatcacca cgggcttaaa ctagctaata ttttattaat 300  
 tgtattctat cctcacatac attctatccc tttttcctca agtcatcctt ctaaactgca 360  
 catctgatca catattgaatc ttagctcctt tacttgcttt ctggccttgg gcagttgttt 420  
 ataatgctct gtgtcctcca ttccctcctgc ctccactagt ggttcatggc ttaatatatg 480  
 taaactatgg cattacctta ctgcttaaaa ctcttaaat taaaaaaaaa aaaaaaaaaa 540  
 tcgagggggg gncc 554

<210> 13

<211> 1106  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> SITE  
 <222> (1017)  
 <223> n equals a,t,g, or c

<400> 13  
 gagcaagctc attttttttt cctatgaggc ttttgtaagt cctgacctgt atttactgtt 60  
 aacttcttag cttgggttca tgcaccccca gtcagtataa ctgtggacct catacccact 120  
 ttggcacagg cttggagtat ggatttatta caggctctgtt tctttttgtt tttctcccat 180  
 ttatggtcct ggacagaagg taagcttcct tgcaacttcc ctgggtccggt gggtagagtt 240  
 ttcttgtccc ctttccagat gttaggtttt aaacaatgac tgttctttct ccatcatgta 300  
 gaccaaaaggc caagtctctgt gtcccatgg gagattaaaa cccaagcccc tatgtctagg 360  
 tccagtgtccc actgatttct ctaattgtga gtctttctgc ttacctagta cctagagttt 420  
 ctcttcccaa gttttaaaaa tatcagttct aagtaggcct agcgtttcta catattttta 480  
 gggagagggg accctttctg tggcagctca gtgttcagca ttcctgtaag ttagcatgct 540  
 ctgtgtatag cagatatcac tagtaatagc atttrgtaag tgatgttcac acatgctgct 600  
 gtcataaaca ctatctcatg ttgtgtaaca ctttcatttt tccaagaact ttataatcag 660  
 ccgacttgaa actcacagtc gtccctcag aaaggcaggg caaatgttgt tatttccaat 720  
 ttgtcagaag ctcagaaagc ttattctgtt gctgacagtc cttgcaaggg tcagaatcag 780  
 gaccggagcc ccagatgcgc tgggtgctact gatgtcccgt gccgggcatg agcccttctg 840  
 tgcaaggagc tccagtgtct cccggacagt gatgatgtga aaacatttag aaccgacct 900  
 cacaataagg cagattttca ttctgtacct aaaacaggaa cacagattta atgcagagca 960  
 aaagggcttt aatcaacaga tatgttcatt ttacacgtag acctatttta caagctnact 1020  
 tgtaagccag aaaatgacat tgcagatttt caagtgagaa caaatgattt ggtccaataa 1080  
 ttaaaaaaaaa aaaaaaaaaa ctcgag 1106

<210> 14  
 <211> 568  
 <212> DNA  
 <213> Homo sapiens

<400> 14  
 gtggatccaa agaattcgca cgagtgccga tcagctcgga ccgaaaaaag tggtttwatt 60  
 cgggctggct tgttgcggtg tgagcggctt gttttatgcc atggcttttt ggttactagg 120  
 tctgccgttg ctgagtttaa ttctgctgtg cattggcagg gtgtttctcg gcgtcggcga 180  
 aagctttgcc agtacggggt ctaccctatg ggggattggc ctgggtggggc cgttgcatat 240  
 cgcccggtt atctcatgga atgggggtggc gacttacggt gcgatggctg ccggggcacc 300  
 gctcgggtgt tacctcaatc agcactgggg gttggctggg gtggcgcggt tgatcgtgtt 360  
 ggcgggtggc gtttcgctgt ggctggcgag tgcgaacca acgtgacgat cgccgcccgt 420  
 aagcgtattg ccttttagcgc atgttggggc gtatttggac ttacggtctg ggacttgcaa 480  
 tgggtaccgt ggggttttggc ggcacgagag tacttctaga gcggccgcgg gcccatcgat 540  
 tttccacccg ggtgggggtac caggtaat 568

<210> 15  
 <211> 3692  
 <212> DNA  
 <213> Homo sapiens

<220>

<221> SITE  
 <222> (518)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (606)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (639)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (2303)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (2441)  
 <223> n equals a,t,g, or c

<400> 15

aattcggcac	aggttgtgtt	tctmatgttc	caggtccggc	caggctggca	gtcctgctg	60
gtcatgtttt	cctcatgtgc	tgtttccaac	cagctcttgg	tctggtaccc	agcaactgcc	120
ttagcagaca	acaaacctgt	agcacctgac	cgacgaatca	gtgggcatgt	gggcatcatc	180
ttcagcatgt	catacctgga	aagcaaggga	ttgctggcta	cagyttcaga	agaccgaagc	240
gttcgtatct	ggaaggtggg	cgacctgcga	gtgcctgggg	gtcgggtgca	gaatattggg	300
cactgctttg	ggcacagcgc	ccgtgtgtgg	caggtcaagc	ttctagagaa	ttaccttata	360
agtgcaggag	aggattgtgt	ctgcttgggtg	tggagccatg	aaggtgagat	tctccaggcc	420
tttcggggac	accagggayg	tggkayccgg	gccatagctg	cccatgagag	gcaggcctgg	480
gtgatcactg	gggtgatga	ctccaggcat	cggctgtngc	acttggtagg	gcgtgggtac	540
cggggattgg	gggtctcggc	tctctgcttc	aagtcccgtg	gtaggccagg	tacactcaag	600
gctgknactc	tggctggctc	ttggcgactg	ctggcagtna	ctgatacagg	ggccctgtat	660
ctctatgacg	tcgaggtcaa	gtgctgggag	cagctgctag	aggataaaca	tttccagtcc	720
tactgcctgc	tggaggcagc	tcctgggtccc	gagggcttcg	gattgtgtgc	tatggccaat	780
ggggaaggtc	gtgtcaaggt	tgtccccatc	aacactccaa	ctgctgctgt	ggaccagacc	840
ctgtttcctg	ggaaggtgca	cagcttgagc	tgggccctgc	gtggttatga	ggagctcctg	900
ttgctggcat	cgggccctgg	cggggtagta	gcttgccatg	agatctcagc	cgcaccctct	960
ggcaaggcca	tctttgtcaa	ggaacgttgt	cggtagctgc	tgcccccaag	caagcagaga	1020
tggcacacat	gcagtgcctt	cctacccccca	ggtracttcc	tgggtgtgtg	tgaccgccgg	1080
ggctctgtgc	tgctattccc	ctccaracca	ggctgtctca	aggaccctgg	ggtgggaggc	1140
aaggctcggg	ctggtgctgg	ggcactgtag	tgggtagtgg	tagtagtggg	ggtgggaatg	1200
ctttcactgg	gttgggcccc	gtgtctaccc	tgccctctct	gcacgggaag	cagggtgtga	1260
cctcagtcac	atgccatggt	ggctatgtgt	ataccacagg	gcgtratgga	gcctactacc	1320
agctgtttgt	acgagacggc	cagctccagc	cagtcctaag	gcagaagtcc	tgtcgaggca	1380
tgaactggct	agctgggctc	cgtatagtgc	ccgatgggag	catggttatc	ctgggtttcc	1440
atgccaatga	gtttgtgggtg	tggaaacctc	ggtcacacga	gaagctgcac	atcgtcaact	1500
gtggtggagg	gcaccgttcg	tgggcattct	ctgatactga	ggcggccatg	gcctttgtct	1560
acctcaagga	tggggatgtc	atgctgtaca	gggctctggg	tggctgcacc	cggccacacg	1620
tgattctccg	ggagggctctg	catggccgtg	agatcacttg	tgtaaagcgt	gtgggcacca	1680
ttacctggg	gcctgaatat	ggagtgccea	gcttcatgca	gcctgatgac	ctggagcctg	1740
gcagtgaggg	gcccgaactg	actgacattg	tgatcacatg	tagtgaggac	actactgtct	1800

gtgtcctagc	actccctaca	accacaggct	cagcccacgc	actcacagct	gtttgtaacc	1860
atatctcctc	ggtacgtgct	gtggctgtgt	ggggcattgg	caccccaggt	ggccctcagg	1920
atcctcagcc	aggcctgact	gcccattgtg	tgtctgcggg	ggggcgggct	gagatgcact	1980
gcttcagcat	catggttact	ccggacccca	gcaccccaag	ccgcctcgcc	tgccatgtca	2040
tgcaccttts	gtcccaccgg	ctagatgagt	attgggaccg	gcaacgcaat	cggcatcgga	2100
tggttaaggt	agaccagag	accaggtaat	atatgtcctc	gggcagggtg	tggatgggt	2160
catgcagatg	ctcccaggct	tgcaggctcc	acctgacagc	tgcatgttgt	ctctgcagggt	2220
acatgtccct	tgctgtgtgt	gaacttgacc	agcccggcct	tggcccccct	gtggctgcag	2280
cctgtagtga	tggggccgta	agntctttct	tttgcaggat	tctgggcgga	ttctgcagct	2340
ccttgctgaa	accttccacc	ataagcratg	tgtcctcaag	gtccactcct	ttacacacga	2400
ggcaccacac	cagaggcgga	ggctcctcct	gtgcagcgca	ntactgatgg	cagcctggct	2460
ttctgggata	tcaccacat	gctagaccat	gactccactg	tcctggagcc	tccagtggat	2520
cctgggcttc	cctaccggct	tggcaccccc	tccttgactc	tccaggccca	cagctgtggt	2580
atcaacagcc	tgcacacctt	gcccaccctg	gagggccacc	atctcgtggc	cagtggcagt	2640
gaagatggat	ccctccatgt	cttcgtgctt	gctgtggaga	tgctacagct	agaagaggct	2700
gtgggagagg	ctgggctggg	accccagctg	cgtgtgctag	aggaatactc	tgtcccctgt	2760
gcacatgctg	cccatgtgac	aggcctcaag	atcctaagcc	caagcatcat	ggtctcagcc	2820
tccattgatc	aacggctgac	cttctggcgt	ctggggcatg	gtgaaccac	cttcatgaat	2880
agcactgtgt	tccatgtgcc	tgatgtggct	gacatggact	gctggcctgt	gagccctgag	2940
tttggccacc	gttgtgccct	tgggggtcag	gggcttgagg	tttacaactg	gtatgactga	3000
ggtatcctgc	ggtggctggc	gtgctgggca	tggggcctgc	tcacagacag	catggagcag	3060
ggaagggctg	tctgtgccc	tgctcagcat	gccttgagg	gaggagggtg	tggccgtggg	3120
ttcctgatgt	cgggtgcagga	gctgaagggt	agtggagtgc	tgccaagaat	atgcccgact	3180
ccccatgaca	agacagaact	ttgtaacaaa	cagtaccaat	ttattttggc	cgtgggtttt	3240
tgcttttttt	ccagttgatg	actttgtgaa	cattcccagg	tattggagcc	tctgtggcct	3300
taaatgtggc	tcagtggagg	gagaccacgc	atagccaggc	cagtatggag	cacctcacgc	3360
acagctctca	gaagctgcag	gcggacgaac	atctgaccaa	agagggtgtg	tcgaggctcc	3420
tgaaagagaa	agggcctgct	ggtctcatcc	tctgcttcc	ttgcctttac	cctataacctc	3480
tctgcacgtc	ccaccccggt	ttgctgtgtg	ctcaccacca	ggatgtgtac	ccggtttag	3540
taggagctga	aatccatgct	gagctgtacc	aggaacttgc	atatctagag	acagagactg	3600
agtcactggc	ccatctcttt	gctcttgtgc	cccaggccag	aataaagaat	agagtgtara	3660
gtraaaaaaa	aaaaaaaaaa	aaaaaactcg	ag			3692

&lt;210&gt; 16

&lt;211&gt; 1428

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 16

agcagggttt	gagcctcctg	gagacattga	atctgaggat	tacactcagc	caatgaagcg	60
cactgtgtca	gataacagcc	tttcaaattc	cagaggagaa	ggcaaaccag	acctcaaatt	120
tgggtggcaa	tccaaaggaa	agttatggcc	gttcatcaaa	aaaaataagg	tactgatggg	180
tggcgtgaaa	tgagttttct	aagggtgtga	gattttgact	tgatctttta	gtcttagaaa	240
aactaagatc	ctaaacctgt	agtttcagaa	tgcaaaagaa	gaagctagt	tgctacctta	300
tgttgagaca	gtattttctt	ttgggtgggt	tatctttgcc	atggccctgt	gtcttatttc	360
agatgcatta	tcctcgtacc	gtgactccca	cactaacaga	gtactgacct	ctccaccgtt	420
tcgcctcatg	cctttccctc	cttcctctcc	tagactgctg	gttaccttgg	ctgggagaga	480
ggatgtagtg	ggacattcct	gtaacacttt	atccgcacat	ctactggaaa	tcgttaccat	540
gttaataact	tggttttgaa	ttcatgttaa	catgtgtacc	catgaacatt	tttcattttc	600
ttttcatagt	gcgatacata	ggtgcatgac	agcattaacc	tggggacgta	gaatatgata	660
aaggcagcat	tactgcttta	actttagaat	gacttactat	ttattaattt	aaacagactg	720
ctgtttccac	aaccttagca	ttgaaggctc	ttcattttct	cccatcaagc	tatgttagtt	780
taggtaatgt	agaaatattt	accctctggc	ttaaagctgg	ttagagtaac	taactagagc	840
tatagtttgc	atgggaaagt	ctgcacgagc	ttcttgtcag	atatttcttg	ctcttctgtc	900

gcattactta	ctaaacctcc	caactctcat	catattcttc	atttaaccac	ctcctacatg	960
ttttcttttg	gaccatggcc	taaaatttaa	ttgttttgtt	tttacttgcg	ttggatttca	1020
aatattattt	gatgcttatt	tttgttttgt	gtcttcttgt	ttctgatttt	tactctgtca	1080
cggctccatc	tcttacatgt	agcttatgtc	ccttttaaca	tccccccatc	agcctcccc	1140
tccccctcct	gcctctgcct	cacctctgc	tggtcccaac	ggccccagc	ctcccaagca	1200
gcaaaaggaa	cccctctccc	accgcttcaa	cgagttcatg	acctccaaac	ccaaaatcca	1260
ctgcttcagg	agcctaaagc	gtggggtaag	ttctgctccg	gaatcctgtc	tctctggcgt	1320
gctttggttg	catgttttgt	tctgcataac	taatttttgt	tgtgaatgaa	tccatttgtt	1380
tttcccataa	catataaaaa	agttaaaaaa	aaaaaaaaaa	aactcgag		1428

&lt;210&gt; 17

&lt;211&gt; 1489

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (7)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (345)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (549)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1408)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1477)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1488)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 17

ggagganagg	atgatgatga	aggaccgtac	acaccattcg	acacccccctc	gggtaaactg	60
gaaacagtga	aatgggcgtt	cacctggccg	ctgagtttcg	tcttatactt	cactgtaccc	120
aactgcaaca	agccgcgctg	ggagaaatgg	ttcatgggtg	cgtttgcttc	ctccacgctg	180
tggatcgcag	ccttctccta	catgatgggtg	tggatgggtc	caatcattgg	ttacaccctg	240
gggattcctg	acgtcatcat	ggggatcacc	ttcctggctg	ctgggaccag	cgtgcctgac	300
tgcattggcca	gcctcattgt	ggccagacaa	rggatggggg	acatngctgt	gtcaaaactcc	360
attgggagca	acgtgtttga	catcctgatt	ggcctcggtc	tcccctgggc	tctgcagacc	420
ctggctgtgg	attacggatc	ctacatccgg	ctgaatagca	gggggctgat	ctactccgta	480

ggcttgcctc	tggcctctgt	ttttgtcacg	gtgttcggcg	tccacctgaa	caagtggcag	540
ctggacaana	agctgggctg	tgggtgcctc	ctcctgtatg	gtgtgttcct	gtgcttctcc	600
atcatgactg	agttcaacgt	gttcaccttt	gtgaacctgc	ccatgtgcgg	ggaccactga	660
accgccgggt	gccacagar	gtcagctcc	ttcttttctg	tgcaatacga	raccggccg	720
cacccgartc	acacaggccc	ctggggccac	ggcgttcgte	tctcctgtgc	tgctcctcagg	780
cctccgctcc	tgttttggtg	gcccargctc	tcccctgccc	catcctcgct	ccccacctc	840
cttgggtcat	gcccaccac	cctttcctgc	ctcctccgtg	tkaagacatc	caacatccac	900
gtgacttttc	cagctccatt	tttgaacagt	gactgagatt	ctagaaaaac	ccggctgcta	960
actggcctga	gccaggcaac	actgattcca	atccctyytc	cttttttaag	ttatttgatg	1020
gaagactcac	ctaatttgtg	acctgagact	gttgaagaaa	tagagaggag	ggggcccgtt	1080
gattacagag	agcatttggg	attttgtttg	gtttggagat	gatgcctagg	ttactgggtt	1140
tggggggatt	gttttctttt	gggggccttc	cccttttact	ccttttcttc	cagagatcaa	1200
gagcttctct	tgcatcttct	tccactgggc	tctggattaa	tcaattaccc	aaaggctgca	1260
cctgccgtgt	tgctgggct	tgcatcccag	atgtgttggg	gtatgcatgg	atgtagtgt	1320
tttttagagga	gccactgggc	aaggccacca	agaacaaatg	catgacattt	tatagccaag	1380
gacgcctcac	taaagtctta	tgggcgtnc	ctggggttgg	gggggcacaa	ggttttgag	1440
gaagaagaca	acttctcat	tccatcatca	ccatctnttt	ctcactang		1489

&lt;210&gt; 18

&lt;211&gt; 1940

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 18

acgcgtccgc	ttcccagaaa	atagatgaca	tcagtgcctc	ttgctacttt	ctcagtcctc	60
actattgctt	tgagggccca	ggtactgaaa	ctgggtgtct	tgagttttgt	gtcagctttt	120
tctccagctc	attatcccc	tccttgcctt	ctgaagcagt	ctagggtaaa	ctagccaggc	180
aggtagttgt	ggaactggtg	attttcaaaa	gccccacttt	agagatcagg	ccacagcttt	240
ttatatcgca	caggacacat	cagcctgagc	tgctgcctca	tgctgtttc	cccaggaaacc	300
tcaactcctt	ggtagaacct	tgggatttta	gaaattgttg	ctttccataa	ctcattttact	360
ccaacagttg	aagttacaca	cattgctccc	aaatttggaa	atagaccaca	gtaccttacc	420
tttcattccc	catctggcct	ttaccttctt	tgttccagt	gttgaaaaca	gttgccatat	480
tcaaagtata	gtagatttca	acctcacaca	aatgacaagt	cccattttac	aatcctagga	540
aggcccacca	atttcatttc	acgcgccagg	gcggctgcag	ttggaggccg	agggcagccc	600
tctgtcact	gaatgtcttg	catgtgctga	ctgctgccc	cagtgtgaa	catgccccac	660
cgccaggcc	cagcactgct	tgttgggtca	gcactagt	ctgctgtcac	atctttgtct	720
gcacagccag	taggattgcc	tcagccagg	ggtttatcag	aagggtgtgca	aggcctttgg	780
gggaactgag	cccctatagt	gggcagtctc	ctttaccttc	ccacctccct	gaaaagcaca	840
gaagacagt	ccttggtttg	tgttttgaag	caaacaagtc	agctttctgg	ctttgccccca	900
aaactgtgat	ggaacataat	aaaactggag	atatggtttt	taacactgca	aaaaggaaaa	960
agcatcaagt	ttctacttct	ggctggaaa	caaaaccaat	ctcagctgac	aaggctgggc	1020
aaactaagtt	ttctgagcc	cattttcctt	tgagccctga	cctagcctgg	ccttacctca	1080
ttaagggtttg	gttaaagcag	tggaaaggag	gaggaggcag	gggtggatgg	gggtgtgggg	1140
aggggatgag	cactctgcag	ccgattaatc	tgttggtagg	ggcccagctt	cttgggagtg	1200
cttattcagc	ccaagagtgg	aggctgttta	cagcgagccc	tggagatggc	agcttgtctc	1260
cagctgggga	ggggtcaggc	ccctaaattg	aagaccactt	tggtagcaga	actgtaggga	1320
ctgggtgagtc	aactcacaga	ttctgcagca	gctgtccac	ccacaataaa	gcaaacgccg	1380
acaggctaga	ccccagattg	caggggctgc	cactacaagg	tgggaccaca	ggctgcctca	1440
ccgggattgt	ttgccactaa	atagctggag	tcacagattg	agataaatgc	cacctcaag	1500
gttgagctga	aaagcataat	cctatgtgat	gaatttatat	gtgttatattt	ttaaaaagct	1560
attttattac	tgcatgttcc	cgctccgtct	tgtgaatgtg	agtecccgcc	accacgtgag	1620
gtgcagtcgt	tgacagcgct	ggtgcaggag	tgccactggc	gcgtgtgtga	tagcatctcg	1680
taggtgttgc	tgacacaagag	ttaaccagag	tcaatgccaa	acacatagta	tgagaagtgt	1740
actttttaag	aaattaattt	atttgagttc	aaatattttt	gaaatataaa	aattgggtgtg	1800

atTTTTttaa	gctataattc	ttgtagacat	tctgtgggta	aaaatttgat	tgtgcttatt	1860
aaaaatggtc	atctatgttt	tgcaattcag	ctacgtgaaa	ataaaatttc	tttgggaagg	1920
tgaaaaaaaa	aaaaaaaaaa					1940

&lt;210&gt; 19

&lt;211&gt; 1592

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 19

ccacgcgtcc	gagcaattta	taaattgata	ccagtaatac	ggtgccttga	caaactagat	60
tgtttgagcc	atgttatgtg	acctcatctt	gttattttaat	ataaaaatgg	caatttatca	120
tctgataata	ctgcagtttt	tctgtagtgt	ttgctctgag	cctgacactg	cactgagtat	180
ttccccactg	taggtcatat	tattgtcccc	atTTTgCCga	tgaagacctg	agaggtgggt	240
aggggcagga	aagtggtcag	tgaggtgctg	gtgggggtgga	ggggccaagg	atcagctgag	300
gctttctgac	ctgagagctg	gcagtgccca	cccaacagct	ctacctgtta	catttctgtt	360
atcctcatgc	catcctctga	ataaacatcc	atcacgctgg	tccttggtag	ccacgacaat	420
gacagcttgt	cagccctggg	ccagtgtgtg	ggcacgtgga	catggggaag	cccagaggtg	480
gcaaggatgg	agtgggttcc	atgtggggag	gataaggggtg	catggagctg	gcaggggagt	540
taagaccggg	aaggccagtg	ggagccagcc	cgggggcagg	ggacctggaa	tgccagcatc	600
agaagacttc	cccgcattgag	gctcttggtg	actggggctc	acctcccttc	tgtctgctgc	660
aggatggggg	gcatgaagtc	caagttcctc	caggctcgag	gcaatacatt	ctcaaaaact	720
gaaaccagcg	ccagcccaca	ctgtcctgtg	tacgtgccgg	atcccacatc	caccatcaag	780
ccggtgagta	ggggagggtc	cagtttccct	gggggctgac	ggatgctgcc	ccaacattgc	840
cctaacagcc	tcctgtcttc	ccaaggttgg	gcaggggtgag	cttgggggtg	aaggagctg	900
actggccacc	aacagtgtcc	tgactcgact	ctccgggacg	cagcgccagc	cagcatgcat	960
aggggcaaag	aagtcattct	tctttcctgc	agcccaggct	tggccctca	gcctgagctc	1020
caccaaacag	gtgtgagtga	gcattctcagg	cctccggcct	gggcagaagc	aaaagcagtg	1080
ctgagaaacg	agtgaacagt	aatgatagca	gctgaagtga	caatgtaatg	gtagtatcag	1140
cagcattgca	aaaatacagt	aacaattgtc	atactaatag	taccaaccat	aattgcgtga	1200
gtataagacc	aagagtagca	gcaggaacaa	cagcaagagt	tgtaaaaaga	gaagcagtca	1260
tgctaggaga	ggtatttacc	aacttctttt	atgtttatac	aaagctgact	atgaaccag	1320
cactgttgta	agccttttacg	gttattaaca	ccctaactct	cacaataatc	ctattgtgtg	1380
acacacaaat	acattcttct	tgaaaataaa	aacacaggcc	aggtgcagtg	gtcacactt	1440
gtactcccag	cattttggga	ggctgaggga	ggaggattac	ttgagcccag	gagttcgaac	1500
tgcatgagct	atgatcacgc	cactgcactc	cagcctgggc	aatacaggga	gaccctgtct	1560
ctttaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aa			1592

&lt;210&gt; 20

&lt;211&gt; 1410

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 20

gcccacgcgt	ccgagaaaaa	tgctgctcag	tttttattgt	ctaccaatgg	taagtataca	60
tattttcttt	ccatgtgccc	actgtgtgta	cctgttgac	atatcctgta	gcctaggaga	120
ggaatcattt	aacagagata	cttgtaaaaa	ggacttttgt	ttttctatac	agaatgtaaa	180
ctctactttt	ttactgtcac	ttgcagtttt	tagattctct	gaaagattct	ctgatagcaa	240
ttttttgttt	actacacctc	caatttgtag	tgaaaagaat	gggctgctat	accattggat	300
ttaggtcagg	tactatttct	gtcatttctc	agtctcgtaa	tcttgggcag	gttactaaca	360
ctgaattgaa	ttttctcag	cagcaacta	gagatagcaa	ttttttatta	tagtattatt	420
atgaatatta	aataacttca	catacatcat	gagtgcaagt	gctcaataaa	tgtaatttta	480
ttctctcttt	ttaagtgttt	gtaaactaca	cagagtatct	caaactgcag	atacaaaata	540

ctcaaaggat	ggtctccatt	ccaggatacg	ctataggaga	gcactttctt	acttgatcac	600
cattagcata	ttgccttctt	cccagcaatc	cacatggctg	gaaggagatt	cctctectac	660
tgtttacttg	ccaagggaaac	atTTTTTgtt	gtTTTTTgag	acaatgtctg	tcgcccaggc	720
tgaagtgcac	tgggtgaatc	acagctcact	gcagcctcga	cctccctacc	tcagtctcct	780
gagtagctgg	gaccacaggt	gagtgccacc	acaccgggct	aattttttta	aaacattttt	840
gtagagcctg	ggtaacatgg	ggtggaacaa	gcctgtagtc	ccagatactc	aggaggctga	900
ggtgaaagga	ttgcttgggc	cagggagggtc	aaggctgcag	tgagccgtga	aaggccactg	960
cactccagcc	tgggtgacag	aatgagacct	tgtctcaaaa	aaaaaaaaaa	agtttcttgg	1020
aacctatacg	tttttttttg	tttttttttt	gaaaagccag	accttgtgcc	cttgttttga	1080
acaccgactg	ggaagatggg	gcttaggttaa	cagccaaacc	tggctgtcag	ctgtgtggga	1140
gccaccaccc	tctctgggaa	gagttcctgc	ttctgtatgg	caagcataaa	tcaagctcag	1200
tctgggttat	ggagaagtgg	aaaattgttt	tgttcctcat	tagtttataa	ttgtatgaaa	1260
tacgatttta	atgaaaactt	ttcagaattc	acgtttgtgt	agatatttca	gagaaccatt	1320
tttactttac	atcctaaaac	tgccttttcc	tatggttttg	tcaataaaac	actatgatgt	1380
tgaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa				1410

&lt;210&gt; 21

&lt;211&gt; 1727

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (979)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1047)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1135)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 21

ccacgcgtcc	ggccatgggt	gccactgtct	gtggcctcct	ggtcttcctg	agcctggggc	60
tggtaccccc	agtcgcgtgc	ctgtttgcac	tcagcgtgcc	caccctgggt	atggagcagg	120
gccgcccggc	gtcctgtgcc	tacagcactg	ccaccctggc	cattgctgtg	gtgccaacag	180
tcctggccaa	cgtgggtgcg	gccgggcagg	tgtctagggtg	tgtcaccgag	ggctccctgg	240
agagtctcct	caataccact	caccagctgc	atgcagcate	cagggctctg	ggccccacag	300
gccaggcagg	cagccggggc	ctgacatttg	aggcccagga	caatggctct	gccttctacc	360
ttcacatgct	cacggtcact	cagcaggctc	tggaggattt	ctctggcctg	gagtccctgg	420
cccgggcagc	agcgttaggg	acccagcgag	tggctcacagg	gctgtttatg	ttgggcctcc	480
tggtggagtc	ggcatggtac	ctccattgct	acctgacaga	cctgcgggtt	gacaatatct	540
acgccactca	acagctgacc	cagcggttgg	cacaggccca	ggctacacac	ctcctggccc	600
ctccacccac	ctggctgctc	caggcggctc	agctgaggct	gtcacaggag	gagctgttga	660
gttgtcttct	aaggctgggg	ctgcttgccc	tgtcctcctg	ggccacggct	gtggcgggtg	720
ccacagacca	tgtagccttc	ctcctggcac	aggctactgt	ggactgggct	cagaagtggc	780
caactgtgcc	catcacgctc	acgggtcaagt	atgatgtggc	atacactgtc	ctgggcttca	840
tccctttcct	cttcaaccag	ctggctccgg	agagccctt	cctctccgtc	cacagctcct	900
accaatggga	gtccgcctc	acctccgccc	gctgcccact	gctacccgcc	cggcgctccc	960
gcgcagctgc	ccgctgggnc	gcggggggcc	tgcagctcct	ggcgggctcc	acgggtgctcc	1020



tggagggcta	cgcccgccgc	ctgcggnatg	ccatcgccgc	ttcctttctt	acagcccagg	1080
aggcgaggag	gatccgccac	ctacacgccc	ggctccagcg	aagacacgac	aggncccaag	1140
gccagcagct	gcccctaggg	gatectttct	gcgtccccac	acccagacct	gcctgcaagc	1200
ctccggcatg	gatagcctac	aggctggatg	ccttaagaac	cgagagcagt	gagggagaag	1260
ggaaagagct	ttggagttgc	agagacctga	gttgtcacct	tggctctgtg	ccgcctccct	1320
gtgtgacctt	gggtaagtca	cttcacctct	ctgagcctcg	gtttctacat	ctgcataacg	1380
acagcatatt	taccattgat	gtgacctact	tcccacgcag	ggatgtgggc	aggatggaag	1440
gaaatactgg	gcatgatagg	cctggataac	cggtaaagaa	ccatgcaaag	gcgaagacaa	1500
ggagtgcaga	gagagctcat	ggttcctcca	ggctgggttg	cgatcaggct	catctcatct	1560
gcaccaactg	ctctacttgt	tagatggaga	ccttgcatca	tgaatttctc	gaaatgctcc	1620
tggaaacttat	ttatatgcct	caaaatcctc	taaactcatt	tatagtaacc	catagtttta	1680
attttataaaa	taaacgtatt	tattaaatct	taaaaaaaaa	aaaaaaa		1727

&lt;210&gt; 22

&lt;211&gt; 1218

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (389)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (740)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1048)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 22

gaaaatagaa	taaatgcccc	tccataagac	taaaattttct	tgtgttttttc	tccttctgag	60
tttaaaatgg	cactggatga	caaatggaaa	gcttgatgct	gctcttaatg	tgccgctagg	120
attccgggga	tttcaaagcc	agtggacggg	agggtggcctc	tgccagtgtc	tgtctggtgt	180
ctgtctgtgt	cactgtgggtg	ctgcctgggc	cacagacctc	ggcaggacct	tgggtgatgg	240
agctcctgtc	tgggtgggtgt	gtgtgggcag	tgctgttctc	gtccacgtta	gaaaagccct	300
cttacttttac	actgagtcac	gctccctctc	caccacggac	cgcagtcccc	ttccctagtg	360
actcgtctgtc	cccttccttt	gttgcgcant	ttctggcttt	aaatgaggag	agcttaagaa	420
tggatgggga	gctcagcact	cacagtaact	gttgggtgaac	tcagggcctg	ctacgtctgg	480
aacacatcaa	gccatttagt	gggtgaggtc	attcactgtt	tttaaatgct	gctgcagctc	540
ttattttctca	tgaagccctt	tatacctatt	aaatacttca	tagtattgaw	taacttagct	600
gsytgctcct	ctctgtcatg	gcaccttttg	ctcatgtgga	ctttawgggtg	cagaaacacg	660
aatcgattgt	cgtaatgaac	aamamccctc	tgaagtggcc	acggcgggta	tgattcgtcc	720
cagttcacgg	gcgagtaacn	gaggtgcgca	gtggcggggc	agctggccca	ggctcgtcag	780
ctgctgtgcg	tgagccagct	cgtcctctgag	tttccttttg	tttgacagca	ttttgtttac	840
agacaccaca	ccaatccttg	gtcttgata	catcagaaaa	gttggagtcc	tagaggtggg	900
tggaggcagg	acttgtaccc	tctccctgca	gcaaagacaa	attcattaag	catttggaac	960
acttgtaag	ttcagtttgt	ctctctctaa	aagttatcac	tagatgactc	tctcattttt	1020
gtgtgtgctg	gttttagatt	tgcctgtnac	ttacgaccag	ggatactggc	tttctattta	1080
tggtagtaat	agcagttctc	cttttaata	aacttatatt	cagccaaaag	agtgattagg	1140
tctatcaaaa	aatgataagg	aaataaacag	tacagatcgt	ctatatttat	ggcaaaaaaa	1200

aaaaaaaaagg gcggccgc

1218

<210> 23

<211> 712

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (26)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (28)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (77)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (117)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (124)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (696)

<223> n equals a,t,g, or c

<400> 23

taggcccggg	acggttacaa	tttacncngg	aaccgctttg	cccataggct	ttgcaaaaag	60
cttttttaggt	gccactntag	aaggtacccc	tgaagggtacc	ggteccggaat	tcccggntgg	120
accnacgcgt	ccgaggaggt	cytttaggaa	gactctcaaa	ggcaaattccc	tgatcccccg	180
ccccaccctt	agccctgccc	tctcaccaga	gcaaaattca	ctgggggactt	ttcccaccac	240
acatggaaat	ctgtccactc	ggaatacctc	tgttttccat	ttcaaattgt	aggggggaggg	300
gatggaacac	ttccagtgat	ggtaagagat	ctggttatgaa	acgaaacacc	ccccgtgtta	360
ataacttggg	ctgaaatctg	tttttatgag	ccggggccccc	tgtgcctcta	gtatacttgt	420
attgactctc	atagttaccc	ttttagtttt	actgtgttct	gtgaaaattt	gtaattgggt	480
gagaatcact	gtgggcgtcc	attcttatcc	aactaaatct	ccacaggttt	tttgagctgg	540
tgtggattag	tttaactctt	gtattcaacc	attagtgtca	ccaccttctc	acattacaat	600
acaattactg	gaagcaagta	ctgcatttcc	tatgcaacaa	aaaaggaaaa	ataaaaaatt	660
gctaattgcta	aaaaaaaaaa	aaaaaaaaaa	aaaaanaaaa	aagggcggcc	gc	712

<210> 24

<211> 1422

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 24

gtctccgctc	ctgtgcccgg	gaagatggtg	ctagggtggtt	gcccgaatca	cgccattttt	60
taacatctct	ttttgatcaa	acaagaaaaa	gcatttgagg	aatgcaaaga	ggactgagaa	120
tactttggct	taaattttgc	ccccagaatc	ttgttggttg	cctactgaag	agatgaaacc	180
atggcagaag	tagaatcctt	atagaaacag	gaccagaaac	acctcccttc	tccaacaaaa	240
ggttcatttt	ggtggctgtc	cgtttgacct	gctgtgcttc	agtttaattg	gcttggaag	300
gggtcagcag	ggtgaaaccg	aaccccagaa	aacttgatga	agaaatgtct	tttgcccgtt	360
ttgattacgt	gcatgcaaac	agcgatttgc	aaagaccgta	tgatgatgat	catgatctta	420
ctgggtgaatt	acagacctga	tgaatttata	gaatgtgaag	acccagtgga	tcatgttgga	480
aatgcaactg	catcccagga	acttggttat	ggttgtctca	agttcggcgg	tcaggcctac	540
agcgacgtgg	aacacacttc	agtccagtgc	catgccttag	atggaattga	gtgtgccagt	600
cctaggacct	ttctacgaga	aaataaacct	tgtataaagt	ataccggaca	ctacttcata	660
accactttac	tctactcctt	cttcctggga	tgttttggtg	tggatcgatt	ctgtttgga	720
cacactggca	ctgcagtagg	gaagctgttg	acgcttgagg	gacttgggat	ttggtggttt	780
gttgacctta	ttttgcta	tactggagg	ctgatgccaa	gtgatggcag	caactgggtc	840
actgtttact	aaaaagagct	gccatcatgg	cccaggagg	cgggtgaaag	ctccgtcttc	900
tgaattcatc	tctacaggct	caaaactcct	ctttgatatc	agacctgatg	ttattttcct	960
tcttttgagg	ggcatttggt	tggttaagaa	ggcttctttg	gactttggaa	tttcaacca	1020
gattttacct	tgcagacgga	atgacaagca	aaaagtgttg	tggggaatca	aatttgttcc	1080
tttcctcatg	cacaaaacat	aaaggatagt	ggcgagttaa	caagctgtgg	atgggtttcc	1140
atagtcttcc	tttctgtaca	ttgctatata	ttcagtcctt	tggagcaagt	ggacctaac	1200
agttgagcaa	aatgaatatt	tggatccatg	ttcctcttgt	gacctgagt	cttcatgcaa	1260
ggagatctga	agctgaacaa	tgaaaatctt	cagcagaaat	agaaatggcc	gtggattgta	1320
atacactctg	aaattctgac	tttctgaatt	taaatgtaga	ataaatttta	ccaacttgga	1380
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaactcg	ag		1422

&lt;210&gt; 25

&lt;211&gt; 1038

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (806)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 25

ggcacgagtg	gctgcagcgg	ggcccgcgtg	gtgcctcctg	aggcggcccc	cggatgaaga	60
gatctgggaa	cccgggagcc	gaggtaacga	acagctcggg	ggcagggcct	gactgctgcg	120
gaggcctcgg	caatattgat	tttagacagg	cagacttctg	cgttatgacc	cggctgctgg	180
gctacgtgga	ccccctggat	cccagctttg	tggctgccgt	catcaccatc	accttcaatc	240
cgctctactg	gaatgtgggt	gcacgatggg	aacacaagac	ccgcaagctg	agcagggcct	300
tcggatcccc	ctacctggcc	tgctactctc	taagcrtcac	catcctgctc	ctgaacttcc	360
tgcgctcgca	ctgcttcacg	caggccatgc	tgagccagcc	caggatggag	agcctggaca	420
ccccgcgggc	ctacagcctg	ggcctcgcg	tcctgggact	gggcgtcgtg	ctcgtgctct	480
ccagcttctt	tgcactgggg	ttcgctggaa	ctttcctagg	tgattacttc	gggatcctca	540
aggaggcgag	agtgaccgtg	ttccccctca	acatcctgga	caacccccatg	tactggggaa	600
gcacagccaa	ctacctgggc	tgggccatca	tgacgcccag	ccccacgggc	ctgctcctga	660
cgggtgctggt	ggccctcacc	tacataragg	ctctcctata	cgaagagccc	ttcaccgctg	720
agatctaccg	gcagaaagcc	tccgggtccc	acaagaggag	ctgattgagc	tgcaacagct	780
ttgctgaagg	cctggccagc	ctcctngctg	ccccaagtgg	caggccctgc	gcagggcgag	840

aatggtgcct	gctgctcagg	gctgcccccg	gcgtgggctg	ccccagtgcc	ttggaacctg	900
ctgccttggg	gaccctggac	gtgccgacat	atggccattg	agctccaacc	cacacattcc	960
cattcaccaa	taaaggcacc	ctgaccccaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1020
aatttggggg	ggggcccc					1038

&lt;210&gt; 26

&lt;211&gt; 1906

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 26

ccgcaacgca	gtgagctcgc	ggccgcgagc	aacaggccgt	gccgrgtttg	catttcctta	60
ctgctttgtc	ttgaagacag	aacgatgcca	aagaaagcaa	agcctacagg	gagtgggaag	120
gaagaggggc	cggctccctg	taagcagatg	aagttagaag	cagctggggg	gccttcagct	180
ttaaactttg	acagtcccag	tagtctcttt	gaaagtttaa	tctcgcccat	caagacagag	240
acttttttca	aggaattctg	ggagcagaag	ccccttctca	ttcagagaga	tgaccctgca	300
ctggccacat	actatgggtc	cctgttcaag	ctaacagatc	tgaagagtct	gtgcagccgg	360
gggatgtact	atggaagaga	tgtgaatgtc	tgccggtgtg	tcaatgggaa	gaagaagggt	420
ttaaataaag	atggcaaagc	acactttctt	cagctgagaa	aagattttga	tcagaaaagg	480
gcaacgattc	agtttcacca	acctcagaga	tttaaggatg	agctttggag	gatccaggag	540
aagctggaat	gttactttgg	ctccttgggt	ggctcgaatg	tgtacataac	tcccgcatat	600
ctcagggcct	gccgccccat	tatgatgatg	tcgagggttt	catcctgcag	ctggagggag	660
agaaacactg	gcgcctctac	cacccactg	tgccctggc	acgagagtac	agcgtggagg	720
ccgaggaaag	gatcggcagg	ccggtgcatg	agtttatgct	gaagccgggt	gatttgttgt	780
actttcccag	aggaaccatt	catcaagcgg	acactcctgc	ggggctggcc	cactcgactc	840
acgtgaccat	cagcacctac	cagaacaatt	catggggaga	tttccttttg	gataccatct	900
cggggcttgt	atttgatact	gcaaaggaag	acgtggagtt	acggaccggc	ataccccggc	960
agctgctcct	gcwgggtgaa	tccacaactg	ttgctacaag	acgattaagt	ggcttcctga	1020
ggacacttgc	agaccggctg	gagggcacca	aagaactgct	ttcctcagac	atgaagaagg	1080
attttattat	gcacagactc	cccccttact	ctgcgggaga	tggggcagag	ctgtcaacac	1140
caggtggaag	gttaccgagg	ctggacagtg	tagtgagact	gcagtttaaa	gaccacattg	1200
tcctcacagt	actgcgggat	caagatcaat	ctgatgaagc	tcaagaaaag	atggtgtaca	1260
tctatcattc	cttaaagaat	agtagagaga	cacacatgat	gggaaatgag	gaggaaacag	1320
agtttcatgg	acttcgcttc	cctttgtcac	atltggatgc	actgaagcaa	atltggaata	1380
gtccagctat	ttctgtcaag	gacctgaaac	ttactacaga	tgaggaaaag	gaaagcctgg	1440
tattatccct	ctggacagaa	tgtttaattc	aagtagtcta	gtgcctttgc	agaatcaaat	1500
gcctactatt	ttatatgcat	atattaaaag	aaaagcaaa	acctgagccg	aggagaggat	1560
gaattcaagt	ttccttacct	gcgtatctac	taacaaacat	gagacctccc	tggtacaggt	1620
ggtcagttgg	ccaaatgtac	taacgggcac	atgaaagaaa	gaacagcaaa	ttaccaagtg	1680
tctcagaaaa	tgacaaaacc	atattttgac	aagtttattt	aatccagtgt	ggtagaaaag	1740
gcacaattcc	aatgtatcat	ttagaattga	atgtcattaa	cctggctttg	ttccttggaa	1800
gaaacaactt	ctttaagag	cttctttggc	tctagaaaaa	tttcaaacia	ttaaaataag	1860
aaaaaatatt	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	ctcgag		1906

&lt;210&gt; 27

&lt;211&gt; 847

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 27

tggtggcgcc	atacatcgcc	ttcacaatgg	cgctctgcag	ctgcgtgttc	tgacgcgtgt	60
cgagcatctt	catctgctcc	atcacgctgt	aaaacacatt	tgaccgcga	gtctgcccgt	120
cctccacggg	ttcattgcgg	cgcagtgtag	acctggggarg	atggscggcc	tgctgctggc	180

tgcttttctg	gctttggtct	cggtgcccag	ggcccaggcc	gtgtggttgg	gaagactgga	240
ccctgagcag	cttcttgggc	cctggtacgt	gcttgcggtg	gcctcccggg	aaaagggtt	300
tgccatggag	aaggacatga	agaacgtcgt	gggggtggtg	gtgacctca	ctccagaaaa	360
caacctgcgg	acgtgtcct	ctcagcacgg	gctgggaggg	tgtgaccaga	gtgtcatgga	420
cctgataaag	cgaaactccg	gatgggtgtt	tgagaatccc	tcaataggcg	tgctggagct	480
ctgggtgctg	gccaccaact	tcagagacta	tgccatcatc	ttcactcagc	tggagttcgg	540
ggacgagccc	ttcaacaccg	tggagctgta	cagtctgacg	gagacagcca	gccaggaggc	600
catggggctc	ttcaccaagt	ggagcaggag	cctgggcttc	ctgtcacagt	agcaggccca	660
gctgcagaag	gacctcacct	gtgctcacia	gatccttctg	tgagtgtctg	gtccccagta	720
gggatggcgc	ccacagggtm	mwgtgacctc	ggccagtgtc	caccacctc	gctcagcggc	780
tcccggggcc	cagcaccagc	tcagaataaa	gcgattccac	agcaaaaaaa	aaaaaaaaaa	840
actcgag						847

&lt;210&gt; 28

&lt;211&gt; 985

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 28

ccacgcgtcc	ggcacagatg	agagcgctcc	gaagactgat	tcagggcagg	atcctgtctcc	60
tgaccatctg	cgctgccggc	attggtggga	cttttcagtt	tggctataac	ctctctatca	120
tcaatgcccc	gaccttgac	attcaggaat	tcaccaatga	gacatggcag	gcgcgtactg	180
gagagccact	gcccgatcac	ctagtccctg	ttatgtggtc	cctcatcgtg	tctctgtatc	240
ccctgggagg	cctctttgga	gcactgcttg	caggtcccctt	ggccatcacg	ctgggaagga	300
agaagtccct	cctggtgaat	aacatctttg	tgggtgtcagc	agcaatcctg	tttggattca	360
gccgcaaagc	aggtcccttt	gagatgatca	tgctgggaag	actgctcgtg	ggagtcaatg	420
caggtgtgag	catgaacatc	cagcccatgt	acctggggga	gagcgccctt	aaggagctcc	480
gaggagctgt	ggccatgagc	tcagccatct	ttacggctct	ggggatcgtg	atgggacagg	540
tggtcggact	cagcactacg	gcggctccgg	ggctccgggg	acttggcagg	ggagctggag	600
gagctggagg	aggagcgcgc	tgcttgccag	ggctgccgtg	cccggcgccc	atgggagctg	660
ttccagcatc	gggccctgag	gagacagggtg	acaagcctcg	tggttctggg	cagtgccatg	720
gagctctgcg	ggaatgactc	ggtgtacgcc	tacgcctcct	ccgtgttccg	gaaggcagga	780
gtgccggaag	cgaagatcca	gtacgcgac	atcgggactg	ggagctgcga	gctgtctcacg	840
gcggttggtt	gtgtgagtct	ggaggggtgcc	cttcctccac	cagccctgtg	gggaggggacc	900
cccaggtcct	ctgcattaaa	ccagtttaca	ctccaaaaaa	aaaaaaaaaa	aaaaaaaaaa	960
aaaaaaaaaa	aaaaaaaaaa	aaaaa				985

&lt;210&gt; 29

&lt;211&gt; 914

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 29

ggcacgagct	aaggctaaga	aagaacactg	tgaatttttc	attatataga	catttttaaaa	60
atactctgat	ctttgtctgtg	ctggcttcta	tagtgtttat	ggggtggaca	actaagacat	120
ttagaattgc	aaaatgccaa	tcagattgga	tggaaacgctg	ggttgacgat	gcatttttga	180
gcttcctttt	ttcgcttatc	cttattgtaa	tcattgtttt	gtggagacca	tcagcaaaaa	240
atcagagata	tgccttcatg	cccttaatatg	atgattctga	tgatgaaatt	gaggaattca	300
tggtaacttc	tgaataattta	accgaaggaa	taaaatttaag	agcctcaaaa	tcagttttcca	360
atggaacagc	taagcctgcc	acttctgaga	actttgatga	agatttgaag	tgggtagaag	420
aaaatattcc	ctcttcattc	acagatgtag	ctcttccagt	gttagtggat	tcagatgagg	480
aaatcatgac	cagatctgaa	atggctgaaa	aaatgttctc	ttcagaaaag	ataatgtgat	540
tggaaccctg	ataagaaatg	tagttaagcc	tgaaggacta	tccttcatca	agactgaaag	600

tgagctttga	tttgatattg	cctaaaaatt	tttattgtgt	tatcttggaa	gtctgtgtat	660
caaaatgaag	aattcagatg	gtaggagggt	ctatagtcct	tttaaagctg	actcttgagt	720
gtcagttgaa	tatccattaa	attggatttg	gaaataacct	gaggaaagta	ttatgataaa	780
gatctgcaca	gatgcctctt	agctgatagg	tggcaggcct	gtgggtttgt	gttctccctc	840
ttttctctgg	aacatatgac	aattccagat	taaagaaaaa	tgttttttta	taaaaaaaaa	900
aaaaaaaaaa	aaaa					914

<210> 30  
 <211> 1183  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (4)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (7)  
 <223> n equals a,t,g, or c

<400> 30	
cacntgnatt catctatcag aacaatgggtg tgagcatgaa gaggcacaga caggctctcca	60
aaatagatgt taggatttgg gtgctacctg acacagaagt aggtctaacc ctccaagtac	120
tggggatgat aggataatca atgaggtata tatatatattg tcatttttga taaaatattg	180
tgaaaattga aggaggacac tcagtaaaca tcctgggact atttgtaagt tatggcaaaa	240
ccagatgaga gaaaaggac agtccccctc gtatcctcgt tgtctcttag taacatcaaa	300
ttgtagttaa aaaaatttta aactatgtac aagctacaaa atagcatctc ttcatggta	360
tgtttgagtg tgtaatttta gtttcttttc tggttgatt tgtggtagtc agatgtgttg	420
gattgattcc aactggacag agtaaggaat tccagcatcc tcttctcgtc tgctcgtgtt	480
acccacaga tcaaacctc aattctagtt ggggatgctg tctagcccca caccatgact	540
gaagccttaa gcaactgtgc gcctccatgt gctttggatc agcaacccca gtggatttct	600
accagagcat tgtgggaaag cagatgtata gtcagggtccc aayagcaaat tgttgggtgt	660
gagagttcta aagtataggg gtgaggggaag agaaggatat gaactcctct gaccttaagc	720
cagcattcat ttaactttta tgtctactta acaagagaac ctggagaaaa ctaccgtatt	780
caagagatta atcaaaatca gtgttttagc caggcgatga cagagaagca ccattcctca	840
ccctccattc ttgtaatgtc tgtaataaat ttcagtgcgt caggatggat gaaccaaga	900
tccagtgaat gattcagctg ttccaagcct tacattttcc atcattcatc atccattctc	960
attcagtgtg acctcttgca ctattgtggg taattttatg taaaaccagt ttatgttttt	1020
ttttttttta tatgtgccta tgtaataaag tctacacact ggctatctct gtagagggtga	1080
ggttttgttt ttagttgttc tactgattat atccttttct gagctatgaa aatgaattat	1140
taataaaaaa tttttgaaca aaaaaaaaaa aaaaaaactc gag	1183

<210> 31  
 <211> 2377  
 <212> DNA  
 <213> Homo sapiens

<400> 31	
aggggcacga gcctaggtgt tgtcgtccct gctagtactc cgggctgtgg gggctcgggtgc	60
ggatattcag tcatgaaatc agggtaggga cttctcccg agcgacgcgg ctggcaagac	120
tgtttgtgtt gcgggggccc gacttcaagg tgattttaca acgagatgct gctctccata	180

gggatgctca	tgctgtcagc	cacacaagtc	tacaccatct	tgactgtcca	gctctttgca	240
ttcttaaacc	tactgcctgt	agaagcagac	atcttagcat	ataactttga	aaatgcacat	300
cagacatttg	atgacctccc	tgcaagattt	ggttatagac	ttccagctga	aggtttaaag	360
ggttttttga	ttaaactcaa	accagagaat	gcctgtgaac	ccatagtgcc	tccaccagta	420
aaagacaatt	catctgggca	ctttcatcgt	gttaattaga	agacttgatt	gtaattttga	480
tataaagggt	ttaaatgcac	agagagcagg	atacaaggca	gccatagttc	acaatgttga	540
ttctgatgac	ctcattagca	tgggatccaa	cgacattgag	gtactaaaga	aaattgacat	600
tccatctgtc	tttattgggt	aatcatcagc	taattctctg	aaagatgaat	tcacatatga	660
aaaagggggc	caccttatct	tagttccaga	atctagtctt	cctttggaat	actaccta	720
tcccttccct	atcatagtgg	gcatctgtct	catcttgata	gtcattttca	tgatcacaaa	780
atctgtccag	gatagacata	gagctagaag	aaacagactt	cgtaaagatc	aacttaagaa	840
acttcctgta	cataaattca	agaaaggaga	tgagtatgat	gtatgtgcca	tttgtttgga	900
tgagtatgaa	gatggagaca	aactcagaat	ccttccctgt	tcccatgctt	atcaytgcaa	960
gtgtgtagac	ccttggctaa	ctaaaaccaa	aaaaacctgt	ccagtgtgca	agcaaaaagt	1020
tgcttcttct	caaggcgatt	cagactctga	cacagacagt	agtcaagaag	aaaatgaagt	1080
gacagaacat	acccctttac	tgagaccttt	agcttctgtc	agtgccagct	catttggggc	1140
tttatcgga	tcccgctcac	atcagaacat	gacagaatct	tcagactatg	aggaagacga	1200
caatgaagat	actgacagta	gtgatgcaga	aatgaaatt	aatgaacatg	atgtcgtggt	1260
ccagttgcag	cctaattggt	aacgggatta	caacatagca	aatactgttt	gactttcaga	1320
agatgattgg	tttatttccc	tttaaaatga	ttaggtatat	actgtaattt	gattttttgc	1380
tcccttcaaa	gatttctgta	gaaataactt	atcttttagt	attctacagt	ttaatcaaat	1440
tactgaaaca	ggacttttga	tctggtattt	atctgccaag	aatatacttc	attcactaat	1500
aatagactgg	tgctgtaact	caagcatcaa	ttcagctctt	cctttggaat	gaaagtatag	1560
ccaaaacata	aaaaaaaaaa	aaatcctcag	tatagcttgc	aattaagacc	tagatcacag	1620
tatttaagtg	ttttgcgttt	tatacatgag	gtcagtgtga	cagccaccta	gcatgaacta	1680
accagcttc	cacctccata	aagttacctt	gagttgttga	gttggaatat	gttctggcat	1740
ttacctgacc	tgccaatcat	tagggagagg	caacaaggta	attcagcctt	tcctcctatc	1800
agcaciaaaga	aactcaaagc	tgttttttcc	ctttctgttc	caaagcagtc	ttatcctgac	1860
aggagcgggc	tatactagtg	cagatttcaa	cacttttttt	taacgtttta	attactatag	1920
tgttatgtag	agatttgatt	gagcagctaa	tgtttctgaa	ctttacttac	taattttcag	1980
tgctcttaag	ggttctgtag	tgttatcaaa	gcaaaaagaa	aatgctgcat	aaaaatacca	2040
aacttcagca	actgttaata	ctcagatcat	atacctctta	ataaatagca	tcttatgcta	2100
attagccctg	ctaaactatg	tacagaggaa	actgttcaag	tattggattt	gaaagtaagt	2160
gacttatgtt	taacagaact	aatgatgtat	tgaaacactg	tattatgaaa	agctaaatta	2220
tacatcattg	taactatgta	gaaagtgtag	actaatgtat	aatcaaaatg	ctaaggattt	2280
ttatatggcc	ttgtatgagg	ggagtttgaa	tggttaataaa	catgttttcc	actttaagat	2340
ccagtaaagt	tctgttctac	tgtagtatta	cttaaaa			2377

&lt;210&gt; 32

&lt;211&gt; 795

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (791)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 32

ggcacagtgc	agcatctacc	taatccaggt	gatctttggt	gctgtggacc	tgcttgccaa	60
gcttgtgggc	ttccttgta	tcaactccct	gggtcgccg	cctgcccaga	tggtgcact	120
gctgctggca	ggcatctgca	tcctgctcaa	tgggtgata	ccccaggacc	agtccattgt	180
ccgaacctct	cttgctgtgc	tggggaagg	ttgtctggct	gcctccttca	actgcatctt	240
cctgtatact	gggaactgta	tcccacaatg	atccggcaga	caggcatggg	aatgggcagc	300

accatggccc	gagtgggcag	catcgtgagc	ccactggtga	gcatgactgc	cgagctctac	360
ccctccatgc	ctctcttcat	ctacggtgct	gttcctgtgg	ccgccagcgc	tgtcactgtc	420
ctctgccag	agaccctggg	ccagccactg	ccagacacgg	tgcaggacct	ggagagcagg	480
aaagggaaac	agacgcgaca	gcaacaagag	caccagaagt	atatgggtccc	actgcaggcc	540
tcagcacaag	agaagaatgg	actctgagga	ctgagaaggg	gccttacaga	accctaaagg	600
gaggggaaggt	cctacaggtc	tccggccacc	cacacaagga	ggaggaagag	gaaatggtga	660
cccaagtgtg	ggggttgtgg	ttcaggaaaag	catcttccca	ggggtccacc	tccctttata	720
aacccccacca	gaaccacatc	attaaaaggt	ttgactgcgm	aaaaaaaaaa	aaaaaaaaaa	780
aactcgaggg	ngggc					795

&lt;210&gt; 33

&lt;211&gt; 2656

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (2652)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 33

gatgagtgcc	tagaagctgc	aatgattgaa	ggagaaattg	agtctttaca	ttcagagaat	60
tcaggaaaat	caggccaaga	gcattggttt	actgaattac	cacctgtgtt	aacatttgaa	120
ttgtcaagat	ttgaatttaa	tcaggcattg	ggaagaccag	aaaaaattca	caacaaatta	180
gaatttcccc	aagttttata	tttggacaga	tacatgcaca	gaaacagaga	aataacaaga	240
attaagaggg	aagagatcaa	gagactgaaa	gattacctca	cggtattaca	acaaaggcta	300
gaaagatatt	taagctatgg	ttccgggtccc	aaacgattcc	ccttggtaga	tgttcttcag	360
tatgcattgg	aatttgcctc	aagtaaacct	gtttgcactt	ctcctgttga	cgatattgac	420
gctagttccc	cacctagtgg	ttccatacca	tcacagacat	taccaagcac	aacagaacaa	480
cagggagccc	tatcttcaga	actgccaagc	acatcacctt	catcagttgc	tgccatttca	540
tcgagatcag	taatacacaa	accatttact	cagtcccgga	tacctccaga	tttggccatg	600
catccggcac	caaggcacat	aacggaggaa	gaactttctg	tgctggaaag	ttgtttacat	660
cgctggagga	cagaaataga	aatgacacc	agagatttgc	aggaaagcat	atccagaatc	720
catcgaacaa	ttgaattaat	gtactctgac	aaatctatga	tacaagttcc	ttatcgatta	780
catgccgttt	tagttcacga	aggccaagct	aatgctgggc	actactgggc	atatattttt	840
gatcatcggtg	aaagcagatg	gatgaagtac	aatgatattg	ctgtgacaaa	atcatcatgg	900
gaagagctag	tgagggactc	ttttggtggt	tatagaaatg	ccagtgcata	ctgtttaatg	960
tacataaatg	ataaggcaca	gttcctaata	caagaggagt	ttaataaaga	aactgggcag	1020
ccccttggtg	gtatagaaac	attaccaccg	gatttgagag	attttggttg	ggaagacaac	1080
caacgatattg	aaaaagaact	agaagaatgg	gatgcacaac	ttgcccagaa	agctttgcag	1140
gaaaagcttt	tagcgtctca	gaaattgaga	gagtcagaga	cttctgtgac	aacagcacaa	1200
gcagcaggag	accagaata	tctagagcag	ccatcaagaa	gtgatttctc	aaagcacttg	1260
aaagaagaaa	ctattcaaat	aattaccaag	gcatcacatg	agcatgaaga	taaaagtcct	1320
gaaacagttt	tgcagtcggc	aattaagttg	gaatatgcaa	ggttggttaa	gttggcccaa	1380
gaagacaccc	caccagaaac	cgattatcgt	ttacatcatg	tagtgggtcta	ctttatccag	1440
aaccaggcac	caaagaaaat	tattgagaaa	acattactag	aacaatttgg	agatagaaat	1500
ttgagttttg	atgaaaggtg	tcacaacata	atgaaagttg	ctcaagccaa	actggaaatg	1560
ataaaacctg	aagaagtaaa	cttgaggagaa	tatgaggagt	ggcatcagga	ttataggaaa	1620
ttcaggggaaa	caactatgta	tctcataatt	gggctagaaa	attttcaaag	agaaagttat	1680
atagattcct	tgctgttcct	catctgtgct	tatcagaata	acaaagaact	cttgtctaaa	1740
ggcttataca	gaggacatga	tgaagaattg	atatcacatt	atagaagaga	atgtttgcta	1800
aaattaaatg	agcaagccgc	agaactcttc	gaatctggag	aggatcgaga	agtaaacaat	1860
ggtttgatta	tcatgaatga	gtttattgtc	ccatttttgc	cattattact	ggtaggatgaa	1920
atggaagaaa	aggatatact	agctgtagaa	gatatgagaa	atcgatgggtg	ttcctacctt	1980



ggtcaagaaa	tggaaccaca	cctccaagaa	aagctgacag	atTTTTtGCC	aaaactgctt	2040
gattgttcta	tgagatttaa	aagtttccat	gagccaccga	agttaccttc	atattccacg	2100
catgaactct	gtgagcgatt	tgcccgaatc	atgttgctcc	tcagtcgaac	tcctgctgat	2160
ggaagataaa	ctgcacactt	tccctgaaca	cactgtataa	actcttttta	gttcttaacc	2220
cttgccctcc	tgtcacaggg	tttgcttggt	gctgctatag	tttttaactt	ttttttattt	2280
taataacygc	aaargacaaa	atgactatac	agactttagt	cagactgcag	acaataaagc	2340
tgaaaatcgc	atggcgctca	gacattttta	ccggaactga	tgtataatca	caaatactaat	2400
tgattttatt	atggcaaaac	tatgcttttg	ccaccttcct	gttgagctat	tactttgctt	2460
ttatcttttc	tttctcaaca	gctttccatt	cagtctggat	ccttccatga	ctacagccat	2520
ttaagtgttc	agcactgtgt	acgatacata	atatttggtg	gcttgtaaat	gaaataaaga	2580
ataaagtttt	atztatggct	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	2640
cgaggggggg	cncaaa					2656

&lt;210&gt; 34

&lt;211&gt; 2566

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (2553)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 34

gcaaatagca	acttcagtag	atcataatat	aaatagaaaa	aaaagatcag	tgcttagatt	60
gttaatgttt	tgTTTTtatt	tgaattattt	tactaaacttg	TTTTtgTTTT	taacctgttc	120
tcgctcagag	tccctctcct	ccccgacagg	accctattca	ggTTTTccct	tcttaaagtc	180
tccccagtg	aggaactctc	tcaacaagg	ccactcctg	gtgcagtact	atagcttttc	240
atcccacctc	agagtcccc	gcaaaaagaa	acaagtgatc	agagtaccag	tcagggtacc	300
tcctaaaagc	ccagcgatgt	cccctccatc	cagtccaagg	tttcactttt	tcaccttttc	360
tggtcccttc	cccaacagct	attaatggta	ttatccattc	aggtctttct	tcacccagg	420
ccttggtgga	ccamccttaa	tcattccagt	gtactgcccc	ctcttaggat	ataccaccam	480
cgstcacaca	ggatctccac	ccagaaacaa	tgacatctgg	ggctcttctc	cagtcacctg	540
gcatgggtatt	tcttacaaac	tttctacctc	ccactggcta	atagctttat	tcaagtasaa	600
ttacacgcca	taaaattttac	tcattttatt	tttttatttt	tattaagtta	ggttgtgttc	660
aggatttact	ctttttaagt	ctgcaattca	cttttttttt	ggtaaattta	gagttgtaca	720
gtcatcacca	tcattccaatt	ttagcacatt	tccatcacct	caaaaagatc	cctcatgccc	780
atTTgtgct	attccacatt	ataaccttcc	acccttgcca	accactaatc	tactttgtgt	840
ctgtatarat	tggttttttc	tgcatatttc	atataaaaaat	ggaacatata	atatttggtc	900
ttaagtattt	ttgaaacata	taattttgtt	gtggaaatag	tagttgattt	tatctatgtc	960
tttatcaggc	ctttctctgt	attgaatttt	cacattgtca	ataccactca	gaaacagtgg	1020
ttyatcctac	tgagcaaggt	tcattgaata	ctgttgccac	tggaatttat	ccctgctgta	1080
accaaagggt	yctycgggtt	gacccacttc	agcttacaaa	gggctgtaaa	rtgagggacc	1140
acatgggtac	mcttcgtgat	caagggtgaag	gsggagattt	gccgtcctgt	cccactgcta	1200
gaatggttga	cgatttgcac	aagtacagag	atgtcattgt	tgtgcctttt	tcaaaagata	1260
cagtttagtga	tggtggggtt	ggcctctgtg	atgaaaagg	tatagaatgt	gatgttttac	1320
tggagccaaa	tacaccatgg	ggccccaaaa	ctggggagct	caatgctttc	ttgtcatatga	1380
aaaactggac	tctacaactg	aaacaacagt	cactgttttc	agaagaagaa	gaatatacca	1440
ctggatctga	ggctactgaa	gatgaagttg	gagatgaaga	agaagtatcc	aagaaacaaa	1500
ggaaaaagga	gaagccaaag	aagttcacta	gacmaccaaa	aaagcaggta	tcttcaccct	1560
gtgcccagag	gaaagaaaag	gcattggaga	aggtaactct	gaattatctg	ktgktaaagt	1620
catatggaaa	aataagcatg	tgagtatagc	cagaaaaaaa	taaaaagagt	aatgaagaca	1680
catggaatgc	tagcaatgta	aaaatgaagt	tttttataga	ctgagattaa	agatctctaa	1740
gatatatattga	caaatgagaa	aaggaaggtg	cagaaacgta	tagtggtata	gtatgctacc	1800

atttgtgtaa	agtagatggg	ggaaatatat	aaataacttc	cttgtatatg	cataaaatgt	1860
ttctggaagg	ctacataaga	actcgataaa	attgggtgcc	tctcaggaag	ggaactgaac	1920
gtgtaaggga	cagaagtgag	agtcttttca	ttatatgtgc	cattatacct	tttgaatttt	1980
aaaccaatat	tatttattca	aaaaattaaa	aatagtcttt	taaattaaaa	ataaatcata	2040
ttttatgata	tttaaaaaa	attcttattt	ctccatgcct	ttgaagggaag	gggtaaaaaa	2100
gccaggtagg	aataagagaa	tagtaataac	caccattggc	taaaagaaaa	actgtgaatt	2160
tcaaaaatgt	gtgatagggt	gagtctgggt	taagatccac	agaattacat	tggacacatt	2220
gtacattcat	ctttgtgtta	agtagcacag	gcatataagt	gggttaattc	taaaaaaaaa	2280
ttgtatcagc	tggctcttgag	cttttgacct	cgtgatctgc	ccgcctcagc	ctcctaaagt	2340
actgggatta	taggcgtgag	ccacaatgcc	tggccacatt	tatgtatttt	tttatattct	2400
gtatcagtta	gcctgtttat	tcacgtaaaa	gttttccacc	atgtcttatt	atccatgggtc	2460
cataggtcat	ctataacaca	tataataaag	tacatcattg	ctgaaaaaaaa	aaaaaaaaaa	2520
actcgagggg	gggtcccgtg	cccaattctc	ctnacatgca	tcgtat		2566

&lt;210&gt; 35

&lt;211&gt; 1668

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 35

aatttcgaac	acccataaaa	ttgtaaagaa	ttgtacagta	cattttaaca	tattkgttg	60
ttacaaycta	tacatttwaw	gttttttaac	cacttcaaag	taagtttcag	acaccaacac	120
atTTTTTaaa	tgatccctac	cattttttta	atgatcccta	ccaaaatgga	aggctgggtat	180
cccaagggtt	tgttccattt	ctcaattcta	gtctgtgaaa	ttgargtctg	atgaccactc	240
ttaaargggc	tgttcattag	ggkgcgggct	gggcattatg	agtgtgtttt	tcattgagkca	300
gtggaaggag	gggcttgttg	tgagcagtg	atgagaaaaa	cggcttggct	ttgcttcttt	360
ttccagctct	gtggccttgg	tcagggttac	tctcttcagt	atcgtaactg	taattgtggag	420
ataaagcctt	cattagttag	gggcacacac	cgcagtattc	cttaagtcac	cttgatgaca	480
agtgaatgca	aggcagctgg	tacctttcag	gtagtagttg	aattcaggta	gtattgttca	540
gttttttttt	ttcccttcat	gttctaagac	cagctgagag	gcaaagttgt	accactgagc	600
tctagtgtgt	gttacctaata	aagsccttgt	tttaaatttc	tgtgatacct	aagaatttca	660
aatctggggt	gtcatggatt	ctttattctt	tttttctccc	ttaaaaagtt	acatttttaga	720
tgaaatcccc	tttyttaaaa	tgggcaaagc	aataattcta	catcatttct	ccccttccct	780
tccacttggt	tagactaaga	tatgttagag	agggaaaggg	tcgttgtttt	agtaaatact	840
attgctgttg	acatgttaat	actattgctg	ttgacatgtt	tactgatggg	ctgtgttcca	900
taattttgtt	ttaggtcttt	tgtttgaaac	agtttactgt	ttttatcagt	tttggtcctt	960
aatttttctt	aacctacagt	ttttctctga	gtacatatgg	tttcattgtt	tgatctactt	1020
tctatctatc	tgaatatgaa	cttctaggat	catgtttatt	ctagtagatg	atgacttaaa	1080
gcctgcagta	taggaggggac	aacgtcaact	actgcagtgt	caataacaag	cttgaaggga	1140
agctaaatgt	ttgttacaaa	tttaagacag	tatttttaag	ccgtttgcat	ttttctaaga	1200
atTTTctata	aagctaattc	tgktattttt	tgtctctaaa	ttaggggaact	gtccagggtt	1260
attgctgccg	ggagactaca	ctgcaaaaata	gataaagtga	atgaaatagt	agaaaccaac	1320
aggtaacttc	atTTTctcaga	ataagggggc	attcctaaat	tttaaaagta	ggkcaactat	1380
tgkcatggaa	taatgtgact	ggtaaaataat	tcattttttc	ttgaatttat	ttatagacct	1440
gatagcaaga	actggcagta	ccaagaaact	atcaagaaag	gagatctgct	actaaacaga	1500
gttcaaaaac	tttccagagt	aattaatatg	taaagccatg	taactaacia	aggatttgct	1560
ttagagataa	ttattttggaa	tttttatagc	ttacttcaca	atgtgccag	gtcagctgta	1620
taaaataaat	actgcattgt	tgttaaaaaa	aaaaaaaaaa	aactcgta		1668

&lt;210&gt; 36

&lt;211&gt; 983

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 36

ccgccccgcct	gccggccccc	gtccggaatt	cccgggtcga	cccacgcgctc	cggggcaagt	60
gagcgagctc	cttcctcacc	gggctgacta	gcctctcctt	tccctgtccc	cctccatcgc	120
tgctctgcag	gaagccagcc	cccagggcca	gtcccggags	ggctgatccg	catctacagc	180
atgaggttct	gcccctattc	tcacaggacc	cgcctcgtcc	tcaaggccaa	agacatcaga	240
catgaagtgg	tcaacattaa	cctgagaaac	aagcctgaat	ggtactatac	aaagcacctt	300
tttgccaca	ttcctgtcct	ggagaccagc	caatgtcaac	tgatctatga	atctgttatt	360
gcttgtgagt	acctggatga	tgcttatcca	ggaaggaagc	tgtttccata	tgacccttat	420
gaacgagctc	gccaaaagat	gttattggag	ctattttgta	aggtcccaca	tttgaccaag	480
gagtgcctgg	tagcgttgag	atgtgggaga	gaatgcacta	atctgaaggc	agccctgcgt	540
caggaattca	gcaacctgga	agagattcct	gagtatcaga	acaccacctt	ctttggtgga	600
acctgtatat	ccatgattga	ttacctcctc	tggccctggg	ttgagcggct	ggatgtgtat	660
gggatactgg	actgtgtgag	ccacackcca	gcctgcggt	ctggatatca	gccatgaagt	720
gggacccccc	agtctgtgct	cttctcatgg	ataagagcat	tttccagggc	ttcttgaatc	780
tctattttca	gaacaaccct	aatgcctttg	actttgggct	gtgctgagtc	tcactgtcca	840
cccttcgct	gtccagaatt	cccagcttg	ttgggagctc	acgtcacggc	ttgtcttggg	900
aaccaatccg	tctctcttcc	ttttctttga	agttcccaat	aaaatgaaaa	caggaaaaaa	960
aaaaaaaaaa	aaagggcggc	cgc				983

&lt;210&gt; 37

&lt;211&gt; 2351

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 37

ccacgcgtcc	ggcagaagca	gcagcagcag	aagacacagc	gccgggtccag	gaggcggctc	60
gagctgttcg	taaagtcgcc	cgacagcttt	ttctccgtag	tatgcgagtt	gacaaaacag	120
ccagagaaca	gggtcccca	ttacaatctt	ttcgagatct	tttcccttgc	taaccggatc	180
tgatttgtgc	gaaaacatgc	cttgcacttg	tacctggagg	aactggagac	agtggattcg	240
acctttagta	gcggctcatct	acctggtgtc	aatagtgggt	gcgggtcccc	tatgcgtgtg	300
ggaattacag	aaactggagg	ttggaataca	caccaaggct	tggtttattg	ctggaatctt	360
tttgctgtga	ctattcctat	atcactgtgg	gtgatattgc	aacacttagt	gcattataca	420
caacctgaac	tacaaaaacc	aataataagg	attctttggg	atggtacctt	tttacagttt	480
tagatagttg	gatagctttg	aaatatcccg	gaattgcaat	atatgtggat	acctgcagag	540
aatgctatga	agcttatgta	atttacaact	ttatgggatt	ccttaccat	tatctaacta	600
accggtatcc	aaatctggta	ttaatccttg	aagccaaaga	tcaacagaaa	catttccctc	660
ctttatgttg	ctgtccacca	tgggctatgg	gagaagtatt	gctgtttagg	tgcaaaactaa	720
gtgtattaca	gtacacagtt	gtcagacctt	tcaccacccat	cgttgcttta	atctgtgagc	780
tgcttggtat	atatgacgaa	gggaacttta	gcttttcaaa	tgcttggact	tatttggtta	840
taataaaca	catgtcacag	ttgtttgcca	tgtattgtct	cctgctcttt	tataaagtac	900
taaaagaaga	actgagccca	atccaacctg	ttggcaaatt	tctttgtgta	aagctggtgg	960
tttttgtttc	tttttgattt	ggcgtttacc	ttttcctaac	atataggcaa	gcagtagtta	1020
ttgctttgtt	ggtaaaagt	ggcgttat	ctgaaaagca	tacgtgggaa	tggcaaaactg	1080
tagaagctgt	ggccaccgga	ctccaggatt	ttattatctg	tattgagatg	ttcctcgctg	1140
ccattgctca	tcattacaca	ttctcatata	aaccatatgt	ccaagaagca	gaagagggtc	1200
catgctttga	ttcctttctt	gccatgtggg	atgtctcaga	tattagagat	gatatttctg	1260
aacaagtaag	gcatgttgga	cggacagtca	ggggacatcc	caggaaaaaa	ttgttcccg	1320
aggatcaaga	tcaaaatgaa	catacaagtt	tattatcatc	atcatcaca	gatgcaattt	1380
ccattgcttc	ttctatgcca	ccttcacca	tgggtcacta	ccaagggttt	ggacacactg	1440
tgactcccca	gactacacct	accacagcta	agatatctga	tgaaatcctt	agtgatacta	1500
taggagagaa	aaaagaacct	tcagataaat	ccgtggattc	ctgaacagta	tggaaaagca	1560
aactgtgcaa	ctactacatt	atatcattac	ctggtatccc	atggattttg	tgcttgggac	1620
agaccataaa	tgatggaaaa	tgtcaacaca	aaaatagctg	aaagccaggt	acaactactg	1680

catttatata	tgtaagtttt	gtatatcaaa	aataattggt	ctaaatttcc	tagacttaga	1740
cttgatttct	taacattagg	gtatcgcata	ctcaaattggt	agacaatgac	cccaactaaa	1800
tcttcttgat	gttacactgc	tttatcaaga	ggatggactt	tttttttttt	gagacagaca	1860
gagtcttgct	ctgtcaccca	ggctggagtg	cagtggcgca	atctcgggtc	actgcaagct	1920
ctgcctccca	agttcatgcc	attctcctgc	ctcagccctc	ccaagtagct	gggactacag	1980
gcacctgcca	ccatgcccgag	ctaatttttt	ttttttcagt	agagacaggg	tctcaccatg	2040
ttagccagga	tggtcttgat	ctgacctcgt	gatccgccga	cctcggcctc	ccaaagtgtc	2100
ggaattacag	gcgtgagcca	ctgcgcctgg	ccaagaatgg	acatttttta	aaaaaacatc	2160
agtacttcct	accactgctg	catgagtata	atgctccgga	attatcagaa	agcataatgc	2220
agaaatacga	attagtggaa	cttaatcatg	tgccatataa	gcttacctaa	caaacagtta	2280
tatccctatt	cctcaactga	atgtctttca	ataaataaga	atttatcatt	taaaaaaaaa	2340
aaaaaaaaaa	a					2351

&lt;210&gt; 38

&lt;211&gt; 1534

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 38

ccacgcgtcc	gccacgcgt	ccggaaatac	taaaaattaa	atgaaaagtt	gtgatgcttg	60
aagtgtctgat	tagtattcgg	actaaagtat	atgaatgaat	aacaattttt	tctctgcaga	120
gactgcagca	tgaaatctca	tgctacattg	actggtgcca	gtggttttta	tttcatagaa	180
ctttcttttc	tggtgttgag	atctgtgctg	ttgggtgctg	ttctgctttg	gcagttccca	240
aagtccctta	caggacaaga	atgatgagtg	gggatataaa	tctcaattcc	agcagctgct	300
cactcacagg	tgtctcgggt	gaagaattgg	gtcttggtga	gcctgtagct	tctctctata	360
tactgtctggg	agatgctgcc	tgtgagtgcc	ttgcttgata	tccaggtgct	aggggctaagg	420
acctctttgt	ggaatagcca	tctttgcttg	aggtctgtgc	aattgtgtat	gcctgcagtg	480
cagtgcctgg	taaggctttc	aaactgtggg	caagaatgta	acaatgcctg	tactcgtga	540
agagacacag	tcgggtgaggt	gagtatggat	tatgccaagg	aaagttttct	gggtcagaga	600
ctttatcctg	ctgcaggaat	taactccatt	gatcaaaaac	agccttaatt	gggatggggc	660
tcgggggcaa	atttcatatg	tgattggcag	gagtctaaac	tgtatagctt	ttctggaggg	720
cattttggca	gtggggatta	aagtgtcaaa	tgtgcatact	ctgtgactgg	acattttcac	780
ttcacagaat	ttatcctaag	gaaagcattg	tacaagtata	cacaaaaggg	tgttcctccg	840
caccataatg	tttagtggtg	ccatcacctg	gggctttatt	aaaaaaggaa	aagttacata	900
aattccagta	aaaccataca	gtggaatatt	ataaagctgc	tgaagaagat	gaagtcaact	960
tctatgtact	attatggaat	gatggtaaag	aaattatgta	caaaagtcac	agatcagcat	1020
gaatagtgtg	atcctatttt	caatatatat	gtgtgtattg	agtgcataatt	atgtaaaggt	1080
ttatatgcat	taattttggg	aggaagaata	tcaaatgcta	atagcgatca	tcaaatgcta	1140
atagtgatta	tgtaaagacc	ctcattttct	acttcctact	tctctgtatt	gtttgaactg	1200
tttataaagg	taaaaccata	gtaatttggg	ctgggtgcgg	tagctcatgc	ctgtaatccc	1260
agcacttttg	gaggccaagt	ggggtggata	tcttgaggtc	agttgtttta	gatgagcctg	1320
accaacatgg	tgaaaccctg	tctctactaa	aaatacaaaa	attggttggg	cttgatgggtg	1380
tgacactgtg	gtcctaacta	cttgggaggg	tgagggtggga	gaattgcttg	aaccaggag	1440
gtggaggtcg	cagtgagctg	agattgcacc	actgcactcc	agcctggatg	atagagcaag	1500
attctctctc	aaaaaaaaata	aaaaaaaaaaa	aaaa			1534

&lt;210&gt; 39

&lt;211&gt; 1182

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 39

agattagagt	gataattctt	gttctttgtg	tattcattta	tacagccctg	ctccatggac	60
------------	------------	------------	------------	------------	------------	----

tactcatgtt	ataataaagg	gatagagaag	ggcatgatga	cgatgtgcgt	tcccagtggtg	120
ctagctgtgg	ctctaccctt	ttttctctca	cttaagaaaa	cttcccagaa	acccgagaag	180
tgagagcatt	ttcccccagg	gaaaaccttg	aattgtgtac	atgtaaatecc	atgggaatct	240
tcagcacttt	attatttagca	tcagattctt	tgttgaactt	aatattattc	ttctttattt	300
tcgctttctc	agtgaagcct	tcttcctcat	cgtttccaag	ttgttgtgtt	tcggtaayck	360
gattatctgt	cattycagag	tcccgtcct	cccackgagc	cacatgcgca	cacacatctc	420
tgtcaggcac	ccctgtcatg	taaggcacgt	tgggtctgcc	agagcggcac	cccttgttcc	480
aactttcagg	tttaatgctt	gagaacattt	gaaggctgtt	gtctggaaaa	gataagtgtt	540
tttatatttc	tttgaatttt	aggagttgtc	taccacaaca	aataaactag	atcacacttt	600
ttaagttcaa	tacttattat	cctcattctg	tggaaaaaat	atattttcta	ttaatcatgt	660
acataatagt	actaattatg	ggccactttg	gctgaacaca	gttttatgct	taggcttaca	720
taattaaggt	tgtaatgtta	tttctggatc	tttgaggcat	tagtagagat	cactgatgaa	780
gtaaactgac	aaacataact	ccttttcttt	ggaaaagatg	gatgctgtct	gctaaactaa	840
tcaagttata	gagccttagg	ccgggtgtgt	cggctcatgc	ctgtggctcc	ggcactttgg	900
gaggccaagg	tgggctggtc	atgaggtcag	gagtttgaga	tcattcctggc	caacatgggtg	960
aaaccccatc	tctactaaac	acacacacac	acacacacac	acacacacac	acacacacac	1020
gccgggcatg	gtggtgggca	tctgtgtgtc	cggctactcg	ggaggctgag	acaggagtgt	1080
cacttcaacc	caagaggcag	aggttgcggt	gagccaagat	catgccattg	cactccagcc	1140
tggaagcctg	ggcaatagag	caasactcca	tctcaaaaaa	aa		1182

&lt;210&gt; 40

&lt;211&gt; 1841

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 40

cgacccacgc	gtccgcacct	gtccctctctg	tgagctctgt	actgtttctcc	gtccctgcaa	60
atagacatga	tgccaccatc	tggaatcatt	gtgtacgtct	ctgctactcc	tcacatcctg	120
ctttgtattt	taatcacttt	catgcttgcc	atcccttcta	ttcataatgg	cagagtttgt	180
gttttattca	tttttttagca	tttggttagca	tttagcacta	atctgtccaa	ataatgaatg	240
ctcaataaac	atgtgtctaa	ttaaactaaa	acaggaggctc	aggctatttc	acctttttcc	300
ccatcacgga	ctgcccttaa	gtctttccct	gaacagaaat	tagcaaattg	aagtaaggaa	360
ccgaggtgtt	agtagcacca	cggactcttc	cactttttca	ccttggcaat	gggaaacatc	420
ctgggggag	agatggcaga	gggagcacat	gggaaccggg	caaagtgtgac	taagagacag	480
cgagtggtga	caaacctcca	cagggtcaca	gatgttggac	atgataaatt	ttgtttcatg	540
aaaaattttg	cttcatgaaa	atgcattatg	cattactttt	acatgaatag	ctaaattgaa	600
cggtagaata	cattgtccca	cttggttaaa	tgtgataaaa	ggagattagt	ggacttgaat	660
ttgtaatcat	ggatgcacac	cacaagggaa	aagcacttgt	tccttctgcc	tcgtcactag	720
tatcagtttg	tggttggtac	ttccaataga	aatgcttcga	aagatgaccc	aagggtcca	780
acaatgacct	tctgaactcc	gttttactga	ctgtttaaaa	taatcctgca	gcttcagatg	840
tattgacttg	gatagaagcc	aacataaate	agacagtgtc	cctgaacaaa	actgaatact	900
tcacactcag	tgccctggtg	cctgtgtgtt	ggagggtattg	gcggcagctt	ctctgtcctt	960
ggtttgtgct	gttttcatgc	agagatagca	acagtaacac	gactaagtga	ccatggctag	1020
ggaaacagcc	tcacattggc	aagtgtgaaa	ggagccaaaa	tatggccagg	catggtggct	1080
cacgcctgta	atcccagcac	tttgggagga	tgagggtgggt	ggatcatttg	aggctcaggag	1140
ttcgagacca	gcctggccaa	catggtgaaa	ccccactctc	actaaaactg	caaaaattgg	1200
ctgggcgtgg	tgggtgggtg	ctgtagtctc	agctactcag	gaggctgaga	caggagaatc	1260
acttgaaccc	gggagatgga	ggttgcagtg	agccaagatt	gcaccactgt	actccagcct	1320
gggtaacaga	gtaagactct	gtctcaaaaa	aaaaaaaaaa	aaaaaaaggg	ggagccaaac	1380
tgtgtttctat	agatgtgcac	ctgagtgtag	gaagaaattt	aatatttagg	gagaaaaatg	1440
ttagatatat	atttttacat	tccttgtgaa	cactggcatt	aatggatagg	gaaccttggt	1500
tttcggggct	ctctgggttt	tggcattgaa	aatctcttgg	ctgggtgcga	tgctcacgcc	1560
tgtaatccca	gcactttggg	aggccgaggt	gggcagatca	tgagttcagg	agttcaagac	1620
cagcctgacc	aacatggtga	aaccccatct	ctactaaaaa	taaaaaaaaaa	attagccagg	1680

catggtggcg	ggcgctgta	atcccagcta	ctcaggaggc	tgaggcagga	gaatcacttg	1740
aaccgggag	gcggaggtg	cagtgaactg	agattgcagc	attgcatccc	agcctgggtg	1800
acagagcgag	actccatctc	aaaaaaaaa	aaaaaaaaa	a		1841

<210> 41  
 <211> 1197  
 <212> DNA  
 <213> Homo sapiens

<400> 41

cccacgcgtc	cgattgggaa	aaggctgtcg	ttaatcactt	ttagcagcag	aaatTTTTTTA	60
TTTTgtgtga	tgctactgtt	ccatgttgaa	gagtcattga	gatgtacaaa	atgtattgac	120
cttatttTgt	actgtgctta	gtgatgtgtc	atatttgcag	caaatacaaa	aaaagTTaag	180
aatgcatgtc	cattgttttg	ctatacatgt	tttatttcat	ttctgctcca	caatttcagc	240
agatgctctt	tcattctgta	tattttgctA	tggaccacag	accctcattg	acatgtattg	300
gaactcctaa	gaccagtgca	gtgctccaag	tatctatgaa	tcaaattggca	gtgttcacat	360
gcttttctcc	cataacttat	aaagcagaag	gtagttttct	ttcccatcac	aatagccatt	420
cttttcttta	tttttcatag	ttattttctt	attaagttat	ctgtaaaaat	aatgcatctc	480
tgtcatctgc	tagcaggcca	ttgttcgagg	ttaaaataca	aattaagaag	aagcaagcaa	540
ataaatcaga	tctggaaacg	aagttagaga	tttttgcaca	aacataatac	ttacaacagt	600
ttcataaaaag	ccaatttatt	atggctactc	ttaacaattc	ctttaaagtt	aaaaactact	660
ataggcgatt	tgtctattat	tcactctttg	tttttataat	tttgttaggt	ttcttttatt	720
caagttactt	tatgtaaatt	acttgagatt	ttatttactg	aaaatcagat	ttcatcattt	780
ctccccagtt	tttctaactg	gctttgattt	ttgtttctta	gcttgattgc	ttgctagttt	840
ttaaatgagg	taaaatatag	attcggtgag	atgcacagat	cttgagtgtg	cagttcaatt	900
gattttgata	aatacaccca	tgtaactgcc	agctgaatca	agataaagac	cattctcatt	960
actccagacc	ttcctgtgtt	tcagtagctt	tggtcacatg	ttcggcagca	tgtgagacga	1020
agttaccta	gccgtaggca	attttatgtg	attctgcata	gtagtcaata	tgggtgataat	1080
gttactttca	tcagaaggct	caaagtaatg	gacctgaaaa	gcaggaaaaa	gaaggggtta	1140
tccagaactt	caagagaact	ctctcaaaga	aagaaaagaa	ggaaaaaaa	aaaaaaa	1197

<210> 42  
 <211> 602  
 <212> DNA  
 <213> Homo sapiens

<400> 42

aattcggcac	agkttgtgtt	tctmatgttc	caggctccggc	caggctggca	gtcctgtctg	60
gtcatgtttt	cctcatgtgc	tgtttccaac	cagctcttgg	tctggtaccc	agcaactgcc	120
ttagcagaca	acaaacctgt	agcacctgac	cgacgaatca	gtgggcatgt	gggcatcatc	180
ttcagcatgt	catacctgga	aagcaaggga	ttgctggcta	cagyttcaga	agaccgaagc	240
gttcgtatct	ggaagggtgg	cgacctgcga	gtgcctgggg	gtcgggtgca	gaatattggg	300
cactgctttg	ggcacagcgc	ccgtgtgtgg	caggctcaagc	ttctagagaa	ttaccttatt	360
agtgcaggag	aggattgtgt	ctgcttgggt	tggagccatg	aagggtgagat	cctccaggcc	420
tttcggggac	accaggatgt	gtacctgggt	gtagtaggag	ctgaaatcca	tgtgtgagctg	480
taccaggaac	ttgcatactt	agagacagag	actgagtcac	tggcccatct	ctttgtctct	540
gtccccaggc	cagaataaag	aatagagtgt	aaaaaaaaa	aaaaaaaaa	aaaaaactcg	600
ag						602

<210> 43  
 <211> 2492  
 <212> DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 43

```

ccacgcgtcc ggaggaagga tgatgatgaa ggaccgtaca caccattcga caccacctcg      60
ggtaaactgg aaacagtgaa atgggcgttc acctggccgc tgagtttcgt cttatacttc      120
actgtaccca actgcaacaa gccgcgctgg gagaaatggt tcatggtgac gtttgcttcc      180
tccacgctgt ggatcgagc cttctcctac atgatggtgt ggatggtcac aatcattggt      240
tacacctggt ggattcctga cgtcatcatg ggggatcacc ttcctggctg ctgggaccag      300
cgtgcctgac tgcattggcca gcctcattgt ggccagacaa gggatggggg acatggctgt      360
gtccaactcc attgggagca acgtgtttga cactcctgatt ggccctcggtc tcccctgggc      420
tctgcagacc ctggctgtgg attacggatc ctacatccgg ctgaatagca gggggctgat      480
ctactccgta ggcttgctcc tggcctctgt ttttgtcaeg gtgttcggcg tccacctgaa      540
caagtggcag ctggacaaga agctgggctg tgggtgcctc ctctctgatg gtgtgttcct      600
gtgtttctcc atcatgactg agttcaacgt gtacaccttt gtgaacctgc ccatgtgcgg      660
ggaccactga gccgcgggt gccacagaa gctcagctcc tcttttctg tgcaatacga      720
gaccggcccg caccctcggt cacacaggcc cctggggcca cggcgctcgt ctctcctgtg      780
ctgtctctag cctccgctc ctgttttggg ggcccaggct cctccctgcc cctcctcgc      840
tccccacctt cctgggttca tggccaccca ccttttctg cctcctcgt gtgaagacat      900
ccaacatcca cgtgactttt ccagctccat ttttgaaacag tgactgagat tctagaaaaa      960
ctggctgcta actggcctga gccaggcaac actgattcca atccctcctc cttttttaag      1020
ttatttgatg gaagactcac ctaatttgtg acctgagact gttgaagaaa tagagaggag      1080
ggggcccggt gattacagag agcatttggg attttgtttg gtttgagat gatgcctagg      1140
ttactgggtt tggggggatt gttttctttt gggggccttc ccttttact cttttcttc      1200
cagagatcaa gagcttctct tgcattctct tccactgggc tctggattaa tcaattacct      1260
aaaggctgca cctgccgtgt tgtctgggct tgcattccag atgtgttga gtatgcatgg      1320
atgtagtgtt ttttagagga gccactgggc aaggccacca agaacaatg catgacattt      1380
tatagccaag gacgcctcac taaagtctta tgggcgtccc ctgggggttg gggggcacia      1440
ggttttggag gaagaagaca acttccctca ttccatcatc accatctctt tctcactagg      1500
ttctttctag ttttcaaagc aataagtcta gcctgccttg gacaaggggg ccccagtta      1560
aacaaactac ccatccatga ggtgccaggc agtcaaaaaa cagaagcttc cccgattgtg      1620
agtccatgag atgtgctctt gttgtaaggc atttgggggtg acagggagtg acccagaggc      1680
caccactgct tttcatgcag gagttacaga cactggtttt cttggaaaat ggagagaagc      1740
gcactttgca cagacgtcgt caattaagtc ccaatttgcc acttggtatt gagtacactg      1800
gaccttgacc actggctttt gggcaaactg ctctctcacg gggcgcttcc gccaaagccg      1860
cccagctgca cccctccctt cctggaggga tggccaggga aggagaaaac agagaactga      1920
cacttttgaa accacagaat gtgtaacatg cagatcgctc aagggcataa gttattgtga      1980
acgtttttgc caatcactgc tcaacagccc tgetagattt tgtatgatgc tgaattatta      2040
tgcagactaa ttccaccag ttgagacaca ccatgcttgt tcaattgtat ttattgaaac      2100
tgtggattct tggcctgtgt gtcccttgta tttactttta gcaactgatc cttatcattc      2160
attcggtatg gttttccctg tcccttgtag acattctggt atgaatttgt aaaaatcccc      2220
tactacaaat tgggtgaatg tttctgtctg tgggtgcgaac cagcattaac ggatggggca      2280
cgtgcccac tggaggaacag gagaagaaat cccaatttg ggctctcaga gctaagacac      2340
acttattgat tctgttcac atttgcact ggtttatggc gattgttttc ttggacggat      2400
agtgtaaaat aaacttctct gttctctaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa      2460
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa

```

&lt;210&gt; 44

&lt;211&gt; 75

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (75)

<223> Xaa equals stop translation

<400> 44

Met Leu His Leu Ala Ala Met Trp Trp Ala Cys Val Thr Thr Leu Val  
1 5 10 15

Phe Thr Leu Val Ser Lys Leu Phe Ile Pro Leu Lys Ser Ser Met Asp  
20 25 30

Gly Glu Met Ser Leu Asp Pro His Ser Cys Val Leu Val Cys Ile Cys  
35 40 45

Phe Pro Leu Arg Phe Val Phe Val Ser Cys Phe Glu Leu Tyr Leu Val  
50 55 60

Gln Ser Ile Val Lys Leu Ser Gln Gln Leu Xaa  
65 70 75

<210> 45

<211> 78

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (78)

<223> Xaa equals stop translation

<400> 45

Met Asp Ala Phe Ala Gly Ser Pro Phe Ser Leu Met Val Pro Lys Cys  
1 5 10 15

Val Leu Ile Leu Phe Cys Leu Val Tyr Ser Leu Gln Cys Ile Gln Pro  
20 25 30

Tyr Ser Ser Leu Leu Asn Ser Ala Ser Leu Pro Tyr His His Gly Leu  
35 40 45

Lys Leu Ala Asn Leu Leu Leu Ile Val Phe Tyr Pro His Ile His Ser  
50 55 60

Ile Pro Phe Ser Ser Ser His Pro Ser Lys Leu His Ile Xaa  
65 70 75

<210> 46

<211> 47

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (47)

<223> Xaa equals stop translation



&lt;400&gt; 46

Met Asp Leu Leu Gln Val Cys Phe Phe Leu Phe Phe Ser His Leu Trp  
 1 5 10 15

Ser Trp Thr Glu Gly Lys Leu Pro Cys Asn Phe Pro Gly Pro Val Gly  
 20 25 30

Arg Val Phe Leu Ser Pro Phe Gln Met Leu Gly Phe Lys Gln Xaa  
 35 40 45

&lt;210&gt; 47

&lt;211&gt; 102

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (102)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 47

Met Ala Phe Trp Phe Thr Gly Leu Pro Leu Leu Ser Leu Ile Leu Leu  
 1 5 10 15

Cys Ile Gly Arg Val Phe Leu Gly Val Gly Glu Ser Phe Ala Ser Thr  
 20 25 30

Gly Ser Thr Leu Trp Gly Ile Gly Leu Val Gly Pro Leu His Thr Ala  
 35 40 45

Arg Val Ile Ser Trp Asn Gly Val Ala Thr Tyr Gly Ala Met Ala Ala  
 50 55 60

Gly Ala Pro Leu Gly Val Tyr Leu Asn Gln His Trp Gly Leu Ala Gly  
 65 70 75 80

Val Ala Ala Leu Ile Val Leu Ala Val Ala Val Ser Leu Trp Leu Ala  
 85 90 95

Ser Ala Asn Pro Thr Xaa  
 100

&lt;210&gt; 48

&lt;211&gt; 382

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (67)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (139)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (141)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (165)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (194)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (344)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (361)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (382)  
 <223> Xaa equals stop translation

<400> 48  
 Met Phe Gln Val Arg Pro Gly Trp Gln Leu Leu Val Met Phe Ser  
 1 5 10 15  
 Ser Cys Ala Val Ser Asn Gln Leu Leu Val Trp Tyr Pro Ala Thr Ala  
 20 25 30  
 Leu Ala Asp Asn Lys Pro Val Ala Pro Asp Arg Arg Ile Ser Gly His  
 35 40 45  
 Val Gly Ile Ile Phe Ser Met Ser Tyr Leu Glu Ser Lys Gly Leu Leu  
 50 55 60  
 Ala Thr Xaa Ser Glu Asp Arg Ser Val Arg Ile Trp Lys Val Gly Asp  
 65 70 75 80  
 Leu Arg Val Pro Gly Gly Arg Val Gln Asn Ile Gly His Cys Phe Gly  
 85 90 95  
 His Ser Ala Arg Val Trp Gln Val Lys Leu Leu Glu Asn Tyr Leu Ile

100	105	110
Ser Ala Gly Glu Asp Cys Val Cys Leu Val Trp Ser His Glu Gly Glu		
115	120	125
Ile Leu Gln Ala Phe Arg Gly His Gln Gly Xaa Gly Xaa Arg Ala Ile		
130	135	140
Ala Ala His Glu Arg Gln Ala Trp Val Ile Thr Gly Gly Asp Asp Ser		
145	150	155
Arg His Arg Leu Xaa His Leu Val Gly Arg Gly Tyr Arg Gly Leu Gly		
	165	170
Val Ser Ala Leu Cys Phe Lys Ser Arg Ser Arg Pro Gly Thr Leu Lys		
	180	185
Ala Xaa Thr Leu Ala Gly Ser Trp Arg Leu Leu Ala Val Thr Asp Thr		
	195	200
Gly Ala Leu Tyr Leu Tyr Asp Val Glu Val Lys Cys Trp Glu Gln Leu		
	210	215
Leu Glu Asp Lys His Phe Gln Ser Tyr Cys Leu Leu Glu Ala Ala Pro		
	225	230
Gly Pro Glu Gly Phe Gly Leu Cys Ala Met Ala Asn Gly Glu Gly Arg		
	245	250
Val Lys Val Val Pro Ile Asn Thr Pro Thr Ala Ala Val Asp Gln Thr		
	260	265
Leu Phe Pro Gly Lys Val His Ser Leu Ser Trp Ala Leu Arg Gly Tyr		
	275	280
Glu Glu Leu Leu Leu Leu Ala Ser Gly Pro Gly Gly Val Val Ala Cys		
	290	295
Leu Glu Ile Ser Ala Ala Pro Ser Gly Lys Ala Ile Phe Val Lys Glu		
	305	310
Arg Cys Arg Tyr Leu Leu Pro Pro Ser Lys Gln Arg Trp His Thr Cys		
	325	330
Ser Ala Phe Leu Pro Pro Gly Xaa Phe Leu Val Cys Gly Asp Arg Arg		
	340	345
Gly Ser Val Leu Leu Phe Pro Ser Xaa Pro Gly Leu Leu Lys Asp Pro		
	355	360
Gly Val Gly Gly Lys Ala Arg Ala Gly Ala Gly Ala Leu Xaa		
	370	375

<211> 46  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (46)  
 <223> Xaa equals stop translation

<400> 49  
 Met Gln Lys Lys Lys Leu Val Cys Tyr Leu Met Leu Arg Gln Tyr Phe  
           1                  5                  10                  15  
 Phe Leu Val Val Val Ser Leu Pro Trp Pro Cys Val Leu Phe Gln Met  
                   20                  25                  30  
 His Tyr Pro Arg Thr Val Thr Pro Thr Leu Thr Glu Tyr Xaa  
                   35                  40                  45

<210> 50  
 <211> 168  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (60)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (64)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (132)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 50  
 Met Val Thr Phe Ala Ser Ser Thr Leu Trp Ile Ala Ala Phe Ser Tyr  
           1                  5                  10                  15  
 Met Met Val Trp Met Val Thr Ile Ile Gly Tyr Thr Leu Gly Ile Pro  
                   20                  25                  30  
 Asp Val Ile Met Gly Ile Thr Phe Leu Ala Ala Gly Thr Ser Val Pro  
                   35                  40                  45  
 Asp Cys Met Ala Ser Leu Ile Val Ala Arg Gln Xaa Met Gly Asp Xaa  
           50                  55                  60  
 Ala Val Ser Asn Ser Ile Gly Ser Asn Val Phe Asp Ile Leu Ile Gly  
           65                  70                  75                  80

Leu Gly Leu Pro Trp Ala Leu Gln Thr Leu Ala Val Asp Tyr Gly Ser  
                             85                            90                            95  
 Tyr Ile Arg Leu Asn Ser Arg Gly Leu Ile Tyr Ser Val Gly Leu Leu  
                             100                            105                            110  
 Leu Ala Ser Val Phe Val Thr Val Phe Gly Val His Leu Asn Lys Trp  
                             115                            120                            125  
 Gln Leu Asp Xaa Lys Leu Gly Cys Gly Cys Leu Leu Leu Tyr Gly Val  
                             130                            135                            140  
 Phe Leu Cys Phe Ser Ile Met Thr Glu Phe Asn Val Phe Thr Phe Val  
                             145                            150                            155                            160  
 Asn Leu Pro Met Cys Gly Asp His  
                             165

<210> 51  
 <211> 50  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (50)  
 <223> Xaa equals stop translation

<400> 51  
 Met Thr Ser Val Pro Leu Ala Thr Phe Ser Val Leu Thr Ile Ala Leu  
     1                            5                            10                            15  
 Arg Ala Gln Val Leu Lys Leu Val Val Leu Ser Phe Val Ser Ala Phe  
                             20                            25                            30  
 Ser Pro Val His Tyr Pro Pro Pro Leu Leu Leu Lys Gln Ser Arg Leu  
                             35                            40                            45  
 Asn Xaa  
     50

<210> 52  
 <211> 41  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (41)  
 <223> Xaa equals stop translation

<400> 52

Met Leu Cys Asp Leu Ile Leu Leu Phe Asn Ile Lys Met Ala Ile Tyr  
 1 5 10 15

His Leu Ile Ile Leu Gln Phe Phe Cys Ser Val Cys Ser Glu Pro Asp  
 20 25 30

Thr Ala Leu Ser Ile Ser Pro Leu Xaa  
 35 40

<210> 53  
 <211> 95  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (95)  
 <223> Xaa equals stop translation

<400> 53  
 Met Leu Leu Ser Phe Tyr Cys Leu Pro Met Val Ser Ile His Ile Phe  
 1 5 10 15

Phe Pro Cys Ala His Cys Val Tyr Leu Leu His Ile Ser Cys Ser Leu  
 20 25 30

Gly Glu Glu Ser Phe Asn Arg Asp Thr Cys Lys Lys Asp Phe Cys Phe  
 35 40 45

Ser Ile Gln Asn Val Asn Ser Thr Phe Leu Leu Ser Leu Ala Val Phe  
 50 55 60

Arg Phe Ser Glu Arg Phe Ser Asp Ser Asn Phe Leu Phe Thr Thr Pro  
 65 70 75 80

Pro Ile Cys Ser Glu Lys Asn Gly Leu Leu Tyr His Trp Ile Xaa  
 85 90 95

<210> 54  
 <211> 485  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (322)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (345)  
 <223> Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (374)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (485)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 54

Met	Val	Ala	Thr	Val	Cys	Gly	Leu	Leu	Val	Phe	Leu	Ser	Leu	Gly	Leu
1				5					10					15	

Val	Pro	Pro	Val	Arg	Cys	Leu	Phe	Ala	Leu	Ser	Val	Pro	Thr	Leu	Gly
			20					25					30		

Met	Glu	Gln	Gly	Arg	Arg	Leu	Leu	Leu	Ser	Tyr	Ser	Thr	Ala	Thr	Leu
		35					40					45			

Ala	Ile	Ala	Val	Val	Pro	Asn	Val	Leu	Ala	Asn	Val	Gly	Ala	Ala	Gly
	50					55					60				

Gln	Val	Leu	Arg	Cys	Val	Thr	Glu	Gly	Ser	Leu	Glu	Ser	Leu	Leu	Asn
65					70					75					80

Thr	Thr	His	Gln	Leu	His	Ala	Ala	Ser	Arg	Ala	Leu	Gly	Pro	Thr	Gly
			85						90					95	

Gln	Ala	Gly	Ser	Arg	Gly	Leu	Thr	Phe	Glu	Ala	Gln	Asp	Asn	Gly	Ser
			100					105					110		

Ala	Phe	Tyr	Leu	His	Met	Leu	Thr	Val	Thr	Gln	Gln	Val	Leu	Glu	Asp
		115					120					125			

Phe	Ser	Gly	Leu	Glu	Ser	Leu	Ala	Arg	Ala	Ala	Ala	Leu	Gly	Thr	Gln
		130					135					140			

Arg	Val	Val	Thr	Gly	Leu	Phe	Met	Leu	Gly	Leu	Leu	Val	Glu	Ser	Ala
145					150					155					160

Trp	Tyr	Leu	His	Cys	Tyr	Leu	Thr	Asp	Leu	Arg	Phe	Asp	Asn	Ile	Tyr
				165					170					175	

Ala	Thr	Gln	Gln	Leu	Thr	Gln	Arg	Leu	Ala	Gln	Ala	Gln	Ala	Thr	His
			180					185					190		

Leu	Leu	Ala	Pro	Pro	Pro	Thr	Trp	Leu	Leu	Gln	Ala	Ala	Gln	Leu	Arg
		195					200					205			

Leu	Ser	Gln	Glu	Glu	Leu	Leu	Ser	Cys	Leu	Leu	Arg	Leu	Gly	Leu	Leu
		210				215					220				

Ala	Leu	Leu	Leu	Val	Ala	Thr	Ala	Val	Ala	Val	Ala	Thr	Asp	His	Val
225					230					235					240

Ala	Phe	Leu	Leu	Ala	Gln	Ala	Thr	Val	Asp	Trp	Ala	Gln	Lys	Leu	Pro	
245								250				255				
Thr	Val	Pro	Ile	Thr	Leu	Thr	Val	Lys	Tyr	Asp	Val	Ala	Tyr	Thr	Val	
260								265				270				
Leu	Gly	Phe	Ile	Pro	Phe	Leu	Phe	Asn	Gln	Leu	Ala	Pro	Glu	Ser	Pro	
275								280				285				
Phe	Leu	Ser	Val	His	Ser	Ser	Tyr	Gln	Trp	Glu	Leu	Arg	Leu	Thr	Ser	
290								295				300				
Ala	Arg	Cys	Pro	Leu	Leu	Pro	Ala	Arg	Arg	Pro	Arg	Ala	Ala	Ala	Pro	
305								310				315				
Leu	Xaa	Ala	Gly	Gly	Leu	Gln	Leu	Leu	Ala	Gly	Ser	Thr	Val	Leu	Leu	
				325				330				335				
Glu	Gly	Tyr	Ala	Arg	Arg	Leu	Arg	Xaa	Ala	Ile	Ala	Ala	Ser	Phe	Phe	
				340				345				350				
Thr	Ala	Gln	Glu	Ala	Arg	Arg	Ile	Arg	His	Leu	His	Ala	Arg	Leu	Gln	
				355				360				365				
Arg	Arg	His	Asp	Arg	Xaa	Gln	Gly	Gln	Gln	Leu	Pro	Leu	Gly	Asp	Pro	
370								375				380				
Ser	Cys	Val	Pro	Thr	Pro	Arg	Pro	Ala	Cys	Lys	Pro	Pro	Ala	Trp	Ile	
385								390				395				
Ala	Tyr	Arg	Leu	Asp	Ala	Leu	Arg	Thr	Glu	Ser	Ser	Glu	Gly	Glu	Gly	
				405				410				415				
Lys	Glu	Leu	Trp	Ser	Cys	Arg	Asp	Leu	Ser	Cys	His	Leu	Gly	Pro	Val	
				420				425				430				
Pro	Pro	Pro	Cys	Val	Thr	Leu	Gly	Lys	Ser	Leu	His	Leu	Ser	Glu	Pro	
435								440				445				
Arg	Phe	Leu	His	Leu	His	Asn	Asp	Ser	Ile	Phe	Thr	Ile	Asp	Val	Thr	
450								455				460				
Tyr	Phe	Pro	Arg	Arg	Asp	Val	Val	Arg	Met	Glu	Gly	Asn	Thr	Gly	His	
465								470				475				
Asp	Arg	Pro	Gly	Xaa												
				485												

```
<210> 55
<211> 115
<212> PRT
<213> Homo sapiens
```



&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (115)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 55

Met Pro Ile His Lys Thr Lys Ile Ser Cys Val Phe Leu Leu Leu Ser  
 1 5 10 15

Leu Lys Trp His Trp Met Thr Asn Gly Lys Leu Asp Ala Ala Leu Asn  
 20 25 30

Val Pro Leu Gly Phe Arg Gly Phe Gln Ser Gln Trp Thr Gly Gly Gly  
 35 40 45

Leu Cys Gln Cys Leu Ser Gly Val Cys Leu Cys His Cys Gly Ala Ala  
 50 55 60

Trp Ala Thr Asp Leu Gly Arg Thr Leu Gly Asp Gly Ala Pro Val Trp  
 65 70 75 80

Trp Val Cys Val Gly Ser Ala Val Pro Val His Val Arg Lys Ala Leu  
 85 90 95

Leu Leu Tyr Thr Glu Ser Cys Ser Leu Ser Thr Thr Asp Arg Ser Pro  
 100 105 110

Leu Pro Xaa  
 115

&lt;210&gt; 56

&lt;211&gt; 50

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (50)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 56

Met Ser Arg Ala Pro Cys Ala Ser Ser Ile Leu Val Leu Thr Leu Ile  
 1 5 10 15

Val Thr Leu Leu Val Leu Leu Cys Ser Val Lys Ile Cys Asn Trp Leu  
 20 25 30

Arg Ile Thr Val Gly Val His Ser Tyr Ser Thr Lys Ser Pro Gln Val  
 35 40 45

Phe Xaa  
 50

<210> 57  
 <211> 172  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (172)  
 <223> Xaa equals stop translation

<400> 57  
 Met Lys Lys Cys Leu Leu Pro Val Leu Ile Thr Cys Met Gln Thr Ala  
           1                  5                  10                  15  
 Ile Cys Lys Asp Arg Met Met Met Ile Met Ile Leu Leu Val Asn Tyr  
                   20                  25                  30  
 Arg Pro Asp Glu Phe Ile Glu Cys Glu Asp Pro Val Asp His Val Gly  
                   35                  40                  45  
 Asn Ala Thr Ala Ser Gln Glu Leu Gly Tyr Gly Cys Leu Lys Phe Gly  
           50                  55                  60  
 Gly Gln Ala Tyr Ser Asp Val Glu His Thr Ser Val Gln Cys His Ala  
           65                  70                  75                  80  
 Leu Asp Gly Ile Glu Cys Ala Ser Pro Arg Thr Phe Leu Arg Glu Asn  
                   85                  90                  95  
 Lys Pro Cys Ile Lys Tyr Thr Gly His Tyr Phe Ile Thr Thr Leu Leu  
                   100                  105                  110  
 Tyr Ser Phe Phe Leu Gly Cys Phe Gly Val Asp Arg Phe Cys Leu Gly  
           115                  120                  125  
 His Thr Gly Thr Ala Val Gly Lys Leu Leu Thr Leu Gly Gly Leu Gly  
           130                  135                  140  
 Ile Trp Trp Phe Val Asp Leu Ile Leu Leu Ile Thr Gly Gly Leu Met  
           145                  150                  155                  160  
 Pro Ser Asp Gly Ser Asn Trp Cys Thr Val Tyr Xaa  
                   165                  170

<210> 58  
 <211> 125  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (101)  
 <223> Xaa equals any of the naturally occurring L-amino acids

&lt;400&gt; 58

Met Leu Ser Gln Pro Arg Met Glu Ser Leu Asp Thr Pro Ala Ala Tyr  
 1 5 10 15

Ser Leu Gly Leu Ala Leu Leu Gly Leu Gly Val Val Leu Val Leu Ser  
 20 25 30

Ser Phe Phe Ala Leu Gly Phe Ala Gly Thr Phe Leu Gly Asp Tyr Phe  
 35 40 45

Gly Ile Leu Lys Glu Ala Arg Val Thr Val Phe Pro Phe Asn Ile Leu  
 50 55 60

Asp Asn Pro Met Tyr Trp Gly Ser Thr Ala Asn Tyr Leu Gly Trp Ala  
 65 70 75 80

Ile Met His Ala Ser Pro Thr Gly Leu Leu Leu Thr Val Leu Val Ala  
 85 90 95

Leu Thr Tyr Ile Xaa Ala Leu Leu Tyr Glu Glu Pro Phe Thr Ala Glu  
 100 105 110

Ile Tyr Arg Gln Lys Ala Ser Gly Ser His Lys Arg Ser  
 115 120 125

&lt;210&gt; 59

&lt;211&gt; 311

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (142)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (311)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 59

Met Leu Leu Trp Leu Leu Gly Trp Leu Glu Cys Val His Asn Ser Arg  
 1 5 10 15

Arg Ser Gln Gly Leu Pro Pro His Tyr Asp Asp Val Glu Val Phe Ile  
 20 25 30

Leu Gln Leu Glu Gly Glu Lys His Trp Arg Leu Tyr His Pro Thr Val  
 35 40 45

Pro Leu Ala Arg Glu Tyr Ser Val Glu Ala Glu Glu Arg Ile Gly Arg  
 50 55 60

Pro Val His Glu Phe Met Leu Lys Pro Gly Asp Leu Leu Tyr Phe Pro

65		70		75		80
Arg Gly Thr Ile His Gln Ala Asp Thr Pro Ala Gly Leu Ala His Ser						
	85			90		95
Thr His Val Thr Ile Ser Thr Tyr Gln Asn Asn Ser Trp Gly Asp Phe						
	100			105		110
Leu Leu Asp Thr Ile Ser Gly Leu Val Phe Asp Thr Ala Lys Glu Asp						
	115			120		125
Val Glu Leu Arg Thr Gly Ile Pro Arg Gln Leu Leu Leu Xaa Val Glu						
	130			135		140
Ser Thr Thr Val Ala Thr Arg Arg Leu Ser Gly Phe Leu Arg Thr Leu						
	145			150		155
Ala Asp Arg Leu Glu Gly Thr Lys Glu Leu Leu Ser Ser Asp Met Lys						
	165			170		175
Lys Asp Phe Ile Met His Arg Leu Pro Pro Tyr Ser Ala Gly Asp Gly						
	180			185		190
Ala Glu Leu Ser Thr Pro Gly Gly Lys Leu Pro Arg Leu Asp Ser Val						
	195			200		205
Val Arg Leu Gln Phe Lys Asp His Ile Val Leu Thr Val Leu Pro Asp						
	210			215		220
Gln Asp Gln Ser Asp Glu Ala Gln Glu Lys Met Val Tyr Ile Tyr His						
	225			230		235
Ser Leu Lys Asn Ser Arg Glu Thr His Met Met Gly Asn Glu Glu Glu						
	245			250		255
Thr Glu Phe His Gly Leu Arg Phe Pro Leu Ser His Leu Asp Ala Leu						
	260			265		270
Lys Gln Ile Trp Asn Ser Pro Ala Ile Ser Val Lys Asp Leu Lys Leu						
	275			280		285
Thr Thr Asp Glu Glu Lys Glu Ser Leu Val Leu Ser Leu Trp Thr Glu						
	290			295		300
Cys Leu Ile Gln Val Val Xaa						
	305			310		

&lt;210&gt; 60

&lt;211&gt; 164

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (2)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (164)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 60

Met	Xaa	Gly	Leu	Leu	Leu	Ala	Ala	Phe	Leu	Ala	Leu	Val	Ser	Val	Pro
1				5					10					15	

Arg	Ala	Gln	Ala	Val	Trp	Leu	Gly	Arg	Leu	Asp	Pro	Glu	Gln	Leu	Leu
		20						25					30		

Gly	Pro	Trp	Tyr	Val	Leu	Ala	Val	Ala	Ser	Arg	Glu	Lys	Gly	Phe	Ala
		35					40					45			

Met	Glu	Lys	Asp	Met	Lys	Asn	Val	Val	Gly	Val	Val	Val	Thr	Leu	Thr
	50					55					60				

Pro	Glu	Asn	Asn	Leu	Arg	Thr	Leu	Ser	Ser	Gln	His	Gly	Leu	Gly	Gly
65					70					75					80

Cys	Asp	Gln	Ser	Val	Met	Asp	Leu	Ile	Lys	Arg	Asn	Ser	Gly	Trp	Val
				85					90					95	

Phe	Glu	Asn	Pro	Ser	Ile	Gly	Val	Leu	Glu	Leu	Trp	Val	Leu	Ala	Thr
			100					105					110		

Asn	Phe	Arg	Asp	Tyr	Ala	Ile	Ile	Phe	Thr	Gln	Leu	Glu	Phe	Gly	Asp
		115					120					125			

Glu	Pro	Phe	Asn	Thr	Val	Glu	Leu	Tyr	Ser	Leu	Thr	Glu	Thr	Ala	Ser
	130					135					140				

Gln	Glu	Ala	Met	Gly	Leu	Phe	Thr	Lys	Trp	Ser	Arg	Ser	Leu	Gly	Phe
145					150					155					160

Leu Ser Gln Xaa

&lt;210&gt; 61

&lt;211&gt; 240

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (240)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 61

Met Arg Ala Leu Arg Arg Leu Ile Gln Gly Arg Ile Leu Leu Leu Thr

41

1	5	10	15
Ile Cys Ala Ala Gly Ile Gly Gly Thr Phe Gln Phe Gly Tyr Asn Leu	20	25	30
Ser Ile Ile Asn Ala Pro Thr Leu His Ile Gln Glu Phe Thr Asn Glu	35	40	45
Thr Trp Gln Ala Arg Thr Gly Glu Pro Leu Pro Asp His Leu Val Leu	50	55	60
Leu Met Trp Ser Leu Ile Val Ser Leu Tyr Pro Leu Gly Gly Leu Phe	65	70	75
Gly Ala Leu Leu Ala Gly Pro Leu Ala Ile Thr Leu Gly Arg Lys Lys	85	90	95
Ser Leu Leu Val Asn Asn Ile Phe Val Val Ser Ala Ala Ile Leu Phe	100	105	110
Gly Phe Ser Arg Lys Ala Gly Ser Phe Glu Met Ile Met Leu Gly Arg	115	120	125
Leu Leu Val Gly Val Asn Ala Gly Val Ser Met Asn Ile Gln Pro Met	130	135	140
Tyr Leu Gly Glu Ser Ala Pro Lys Glu Leu Arg Gly Ala Val Ala Met	145	150	155
Ser Ser Ala Ile Phe Thr Ala Leu Gly Ile Val Met Gly Gln Val Val	165	170	175
Gly Leu Ser Thr Thr Ala Ala Pro Gly Leu Arg Gly Leu Gly Arg Gly	180	185	190
Ala Gly Gly Ala Gly Gly Gly Ala Arg Cys Leu Pro Gly Leu Pro Cys	195	200	205
Pro Ala Pro Met Gly Ala Val Pro Ala Ser Gly Pro Glu Glu Thr Gly	210	215	220
Asp Lys Pro Arg Gly Ser Gly Gln Cys His Gly Ala Leu Arg Glu Xaa	225	230	235
			240

&lt;210&gt; 62

&lt;211&gt; 130

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (130)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 62

Met Glu Arg Trp Val Asp Asp Ala Phe Trp Ser Phe Leu Phe Ser Leu  
 1 5 10 15

Ile Leu Ile Val Ile Met Phe Leu Trp Arg Pro Ser Ala Asn Asn Gln  
 20 25 30

Arg Tyr Ala Phe Met Pro Leu Ile Asp Asp Ser Asp Asp Glu Ile Glu  
 35 40 45

Glu Phe Met Val Thr Ser Glu Asn Leu Thr Glu Gly Ile Lys Leu Arg  
 50 55 60

Ala Ser Lys Ser Val Ser Asn Gly Thr Ala Lys Pro Ala Thr Ser Glu  
 65 70 75 80

Asn Phe Asp Glu Asp Leu Lys Trp Val Glu Glu Asn Ile Pro Ser Ser  
 85 90 95

Phe Thr Asp Val Ala Leu Pro Val Leu Val Asp Ser Asp Glu Glu Ile  
 100 105 110

Met Thr Arg Ser Glu Met Ala Glu Lys Met Phe Ser Ser Glu Lys Ile  
 115 120 125

Met Xaa  
 130

&lt;210&gt; 63

&lt;211&gt; 61

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (61)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 63

Met Phe Glu Cys Val Ile Leu Val Ser Phe Leu Val Val Phe Val Val  
 1 5 10 15

Val Arg Cys Val Gly Leu Ile Pro Thr Gly Gln Ser Lys Glu Phe Gln  
 20 25 30

His Pro Leu Pro Ala Cys Ser Cys Tyr Pro Thr Asp Gln Thr Leu Asn  
 35 40 45

Ser Ser Trp Gly Cys Cys Leu Ala Pro His His Asp Xaa  
 50 55 60

<210> 64  
 <211> 98  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (98)  
 <223> Xaa equals stop translation

<400> 64  
 Met Leu Leu Ser Ile Gly Met Leu Met Leu Ser Ala Thr Gln Val Tyr  
           1                  5                  10                  15  
 Thr Ile Leu Thr Val Gln Leu Phe Ala Phe Leu Asn Leu Leu Pro Val  
                   20                  25                  30  
 Glu Ala Asp Ile Leu Ala Tyr Asn Phe Glu Asn Ala Ser Gln Thr Phe  
           35                  40                  45  
 Asp Asp Leu Pro Ala Arg Phe Gly Tyr Arg Leu Pro Ala Glu Gly Leu  
           50                  55                  60  
 Lys Gly Phe Leu Ile Asn Ser Lys Pro Glu Asn Ala Cys Glu Pro Ile  
           65                  70                  75                  80  
 Val Pro Pro Pro Val Lys Asp Asn Ser Ser Gly His Phe His Arg Val  
                   85                  90                  95

Asn Xaa

<210> 65  
 <211> 54  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (54)  
 <223> Xaa equals stop translation

<400> 65  
 Met Ala Ala Leu Leu Leu Ala Gly Ile Cys Ile Leu Leu Asn Gly Val  
           1                  5                  10                  15  
 Ile Pro Gln Asp Gln Ser Ile Val Arg Thr Ser Leu Ala Val Leu Gly  
                   20                  25                  30  
 Lys Gly Cys Leu Ala Ala Ser Phe Asn Cys Ile Phe Leu Tyr Thr Gly  
           35                  40                  45

Asn Cys Ile Pro Gln Xaa



50.

<210> 66  
 <211> 64  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (64)  
 <223> Xaa equals stop translation

<400> 66  
 Met His Trp Asn Leu Pro Gln Val Asn Leu Phe Ala Leu Leu Leu Leu  
           1                  5                  10                  15  
 Thr Ile Leu Thr Leu Val Pro His Leu Val Val Pro Tyr His His Arg  
                   20                  25                  30  
 His Tyr Gln Ala Gln Gln Asn Asn Arg Glu Pro Tyr Leu Gln Asn Cys  
           35                  40                  45  
 Gln Ala His His Leu His Gln Leu Leu Pro Phe His Arg Asp Gln Xaa  
           50                  55                  60

<210> 67  
 <211> 107  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (107)  
 <223> Xaa equals stop translation

<400> 67  
 Met Phe Cys Phe Tyr Leu Asn Tyr Phe Thr Asn Leu Phe Leu Phe Leu  
           1                  5                  10                  15  
 Thr Cys Ser Arg Ser Glu Ser Leu Ser Ser Pro Thr Gly Pro Tyr Ser  
                   20                  25                  30  
 Gly Phe Pro Phe Leu Lys Ser Pro Pro Val Arg Asn Ser Leu Asn Lys  
           35                  40                  45  
 Gly Pro Leu Leu Val Gln Tyr Tyr Ser Phe Ser Ser His Leu Arg Val  
           50                  55                  60  
 Pro Arg Lys Lys Lys Gln Val Ile Arg Val Pro Val Arg Val Pro Pro  
           65                  70                  75                  80

Lys Ser Pro Ala Met Ser Pro Pro Ser Ser Pro Arg Phe His Phe Phe  
85 90 95

Thr Phe Ser Gly Pro Phe Pro Asn Ser Tyr Xaa  
100 105

<210> 68  
<211> 45  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (45)  
<223> Xaa equals stop translation

<400> 68  
Met Arg Lys Thr Ala Trp Leu Cys Phe Phe Phe Gln Leu Cys Gly Leu  
1 5 10 15

Gly Gln Val Thr Ser Leu Gln Tyr Arg Asn Cys Asn Val Glu Ile Lys  
20 25 30

Pro Ser Leu Val Arg Gly Thr His Arg Ser Ile Pro Xaa  
35 40 45

<210> 69  
<211> 43  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (43)  
<223> Xaa equals stop translation

<400> 69  
Met Asn Leu Leu Leu Leu Val Ser Thr Trp Met Met Leu Ile Gln Glu  
1 5 10 15

Gly Ser Cys Phe His Met Thr Leu Met Asn Glu Leu Ala Lys Arg Cys  
20 25 30

Tyr Trp Ser Tyr Phe Val Arg Ser His Ile Xaa  
35 40

<210> 70  
<211> 58  
<212> PRT  
<213> Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (58)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 70

Met	Pro	Cys	Thr	Cys	Thr	Trp	Arg	Asn	Trp	Arg	Gln	Trp	Ile	Arg	Pro
1				5					10					15	

Leu	Val	Ala	Val	Ile	Tyr	Leu	Val	Ser	Ile	Val	Val	Ala	Val	Pro	Leu
			20					25					30		

Cys	Val	Trp	Glu	Leu	Gln	Lys	Leu	Glu	Val	Gly	Ile	His	Thr	Lys	Ala
			35				40					45			

Trp	Phe	Ile	Ala	Gly	Ile	Phe	Leu	Leu	Xaa
	50					55			

&lt;210&gt; 71

&lt;211&gt; 45

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (45)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 71

Met	Lys	Ser	His	Ala	Thr	Leu	Thr	Gly	Gly	Ser	Gly	Phe	Tyr	Phe	Ile
1				5					10					15	

Glu	Leu	Ser	Phe	Leu	Leu	Leu	Arg	Ser	Val	Leu	Leu	Val	Leu	Val	Leu
			20					25					30		

Leu	Trp	Gln	Phe	Pro	Lys	Ser	Leu	Thr	Gly	Gln	Glu	Xaa
			35				40					45

&lt;210&gt; 72

&lt;211&gt; 71

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (43)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (44)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (49)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (52)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (56)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (71)  
 <223> Xaa equals stop translation

<400> 72  
 Met Gly Ile Phe Ser Thr Leu Leu Leu Ala Ser Asp Ser Leu Leu Asn  
           1                  5                  10                  15  
 Leu Ile Leu Phe Phe Phe Ile Phe Ala Phe Ser Val Lys Leu Ser Ser  
                   20                  25                  30  
 Ser Ser Phe Pro Ser Cys Cys Val Ser Val Xaa Xaa Leu Ser Val Ile  
           35                  40                  45  
 Xaa Glu Ser Xaa Ser Ser His Xaa Ala Thr Cys Ala His Thr Ser Leu  
           50                  55                  60  
 Ser Gly Thr Pro Val Met Xaa  
           65                  70

<210> 73  
 <211> 44  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (44)  
 <223> Xaa equals stop translation

<400> 73  
 Met Met Ser Pro Ser Gly Ile Ile Val Tyr Val Ser Ala Thr Pro His  
           1                  5                  10                  15  
 Ile Leu Leu Cys Ile Leu Ile Thr Phe Met Leu Ala Ile Pro Ser Ile  
                   20                  25                  30  
 His Asn Gly Arg Val Cys Val Leu Phe Ile Phe Xaa

35

40

&lt;210&gt; 74

&lt;211&gt; 43

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (43)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 74

Met	His	Val	His	Cys	Phe	Ala	Ile	His	Val	Leu	Phe	His	Phe	Cys	Ser
1				5					10					15	

Thr	Ile	Ser	Ala	Asp	Ala	Leu	Ser	Phe	Cys	Ile	Phe	Cys	Tyr	Gly	Pro
			20					25					30		

Gln	Thr	Leu	Ile	Asp	Met	Tyr	Trp	Asn	Ser	Xaa
		35					40			

&lt;210&gt; 75

&lt;211&gt; 178

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (67)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (178)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 75

Met	Phe	Gln	Val	Arg	Pro	Gly	Trp	Gln	Leu	Leu	Leu	Val	Met	Phe	Ser
1				5					10					15	

Ser	Cys	Ala	Val	Ser	Asn	Gln	Leu	Leu	Val	Trp	Tyr	Pro	Ala	Thr	Ala
			20					25					30		

Leu	Ala	Asp	Asn	Lys	Pro	Val	Ala	Pro	Asp	Arg	Arg	Ile	Ser	Gly	His
		35					40					45			

Val	Gly	Ile	Ile	Phe	Ser	Met	Ser	Tyr	Leu	Glu	Ser	Lys	Gly	Leu	Leu
	50					55					60				

Ala	Thr	Xaa	Ser	Glu	Asp	Arg	Ser	Val	Arg	Ile	Trp	Lys	Val	Gly	Asp
65						70					75				80

Leu Arg Val Pro Gly Gly Arg Val Gln Asn Ile Gly His Cys Phe Gly  
85 90 95

His Ser Ala Arg Val Trp Gln Val Lys Leu Leu Glu Asn Tyr Leu Ile  
100 105 110

Ser Ala Gly Glu Asp Cys Val Cys Leu Val Trp Ser His Glu Gly Glu  
115 120 125

Ile Leu Gln Ala Phe Arg Gly His Gln Asp Val Tyr Pro Val Val Val  
130 135 140

Gly Ala Glu Ile His Ala Glu Leu Tyr Gln Glu Leu Ala Tyr Leu Glu  
145 150 155 160

Thr Glu Thr Glu Ser Leu Ala His Leu Phe Ala Leu Val Pro Arg Pro  
165 170 175

Glu Xaa

<210> 76

<211> 49

<212> PRT

<213> Homo sapiens

**<220>**

<221> SITE

<222> (49)

<223> Xaa equals stop translation

<400> 76

Met Val Thr Phe Ala Ser Ser Thr Leu Trp Ile Ala Ala Phe Ser Tyr  
1 5 10 15

Met Met Val Trp Met Val Thr Ile Ile Gly Tyr Thr Leu Gly Ile Pro  
20 25 30

Asp Val Ile Met Gly Asp His Leu Pro Gly Cys Trp Asp Gln Arg Ala  
35 40 45

**Xaa**

<210> 77

<211> 14

<212> PRT

<213> Homo sapiens

<400> 77

Asn Tyr Phe Pro Val His Thr Val Gln Pro Asn Trp Tyr Val  
1 5 10

<210> 78  
<211> 31  
<212> PRT  
<213> Homo sapiens

<400> 78  
Pro Val Phe Thr Val Asn Phe Leu Ala Trp Val His Ala Pro Pro Val  
1 5 10 15  
Ser Ile Thr Val Asp Leu Ile Pro Thr Leu Ala Gln Ala Trp Ser  
20 25 30

<210> 79  
<211> 33  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (19)  
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 79  
Trp Ile Gln Arg Ile Arg Thr Ser Ala Asp Gln Leu Gly Pro Lys Lys  
1 5 10 15  
Val Val Xaa Phe Gly Leu Ala Cys Cys Gly Val Ser Gly Leu Phe Tyr  
20 25 30

Ala

<210> 80  
<211> 351  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (78)  
<223> Xaa equals any of the naturally occurring L-amino acids

<220>  
<221> SITE  
<222> (326)  
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 80  
Pro Pro Gly Leu Cys Ala Ala Ile Pro Leu Gln Thr Arg Ser Ala Gln  
1 5 10 15  
Gly Pro Trp Gly Gly Arg Gln Gly Ser Gly Trp Cys Trp Gly Thr Val

20					25					30						
Val	Gly	Ser	Gly	Ser	Ser	Gly	Gly	Gly	Asn	Ala	Phe	Thr	Gly	Leu	Gly	
35					40					45						
Pro	Val	Ser	Thr	Leu	Pro	Ser	Leu	His	Gly	Lys	Gln	Gly	Val	Thr	Ser	
50					55					60						
Val	Thr	Cys	His	Gly	Gly	Tyr	Val	Tyr	Thr	Thr	Gly	Arg	Xaa	Gly	Ala	
65					70					75					80	
Tyr	Tyr	Gln	Leu	Phe	Val	Arg	Asp	Gly	Gln	Leu	Gln	Pro	Val	Leu	Arg	
85					90					95						
Gln	Lys	Ser	Cys	Arg	Gly	Met	Asn	Trp	Leu	Ala	Gly	Leu	Arg	Ile	Val	
100					105					110						
Pro	Asp	Gly	Ser	Met	Val	Ile	Leu	Gly	Phe	His	Ala	Asn	Glu	Phe	Val	
115					120					125						
Val	Trp	Asn	Pro	Arg	Ser	His	Glu	Lys	Leu	His	Ile	Val	Asn	Cys	Gly	
130					135					140						
Gly	Gly	His	Arg	Ser	Trp	Ala	Phe	Ser	Asp	Thr	Glu	Ala	Ala	Met	Ala	
145					150					155					160	
Phe	Ala	Tyr	Leu	Lys	Asp	Gly	Asp	Val	Met	Leu	Tyr	Arg	Ala	Leu	Gly	
165					170					175						
Gly	Cys	Thr	Arg	Pro	His	Val	Ile	Leu	Arg	Glu	Gly	Leu	His	Gly	Arg	
180					185					190						
Glu	Ile	Thr	Cys	Val	Lys	Arg	Val	Gly	Thr	Ile	Thr	Leu	Gly	Pro	Glu	
195					200					205						
Tyr	Gly	Val	Pro	Ser	Phe	Met	Gln	Pro	Asp	Asp	Leu	Glu	Pro	Gly	Ser	
210					215					220						
Glu	Gly	Pro	Asp	Leu	Thr	Asp	Ile	Val	Ile	Thr	Cys	Ser	Glu	Asp	Thr	
225					230					235					240	
Thr	Val	Cys	Val	Leu	Ala	Leu	Pro	Thr	Thr	Thr	Gly	Ser	Ala	His	Ala	
245					250					255						
Leu	Thr	Ala	Val	Cys	Asn	His	Ile	Ser	Ser	Val	Arg	Ala	Val	Ala	Val	
260					265					270						
Trp	Gly	Ile	Gly	Thr	Pro	Gly	Gly	Pro	Gln	Asp	Pro	Gln	Pro	Gly	Leu	
275					280					285						
Thr	Ala	His	Val	Val	Ser	Ala	Gly	Gly	Arg	Ala	Glu	Met	His	Cys	Phe	
290					295					300						
Ser	Ile	Met	Val	Thr	Pro	Asp	Pro	Ser	Thr	Pro	Ser	Arg	Leu	Ala	Cys	
305					310					315					320	



His Val Met His Leu Xaa Ser His Arg Leu Asp Glu Tyr Trp Asp Arg  
325 330 335

Gln Arg Asn Arg His Arg Met Val Lys Val Asp Pro Glu Thr Arg  
340 345 350

```
<210> 81
<211> 38
<212> PRT
<213> Homo sapiens
```

<400> 81  
Pro Pro Gly Leu Cys Ala Ala Ile Pro Leu Gln Thr Arg Ser Ala Gln  
1 5 10 15

Gly Pro Trp Gly Gly Arg Gln Gly Ser Gly Trp Cys Trp Gly Thr Val  
20 25 30

Val Gly Ser Gly Ser Ser  
35

```
<210> 82
<211> 40
<212> PRT
<213> Homo sapiens
```

```
<220>
<221> SITE
<222> (40)
<223> Xaa equals any of the naturally occurring L-amino acids
```

<400> 82  
Gly Gly Gly Asn Ala Phe Thr Gly Leu Gly Pro Val Ser Thr Leu Pro  
1 5 10 15

Ser Leu His Gly Lys Gln Gly Val Thr Ser Val Thr Cys His Gly Gly  
20 25 30

Tyr Val Tyr Thr Thr Gly Arg Xaa  
35 40

```
<210> 83
<211> 40
<212> PRT
<213> Homo sapiens
```

<400> 83  
Gly Ala Tyr Tyr Gln Leu Phe Val Arg Asp Gly Gln Leu Gln Pro Val  
1 5 10 15

Leu Arg Gln Lys Ser Cys Arg Gly Met Asn Trp Leu Ala Gly Leu Arg

20 25 30  
 Ile Val Pro Asp Gly Ser Met Val  
 35 40  
  
 <210> 84  
 <211> 41  
 <212> PRT  
 <213> Homo sapiens  
  
 <400> 84  
 Ile Leu Gly Phe His Ala Asn Glu Phe Val Val Trp Asn Pro Arg Ser  
 1 5 10 15  
 His Glu Lys Leu His Ile Val Asn Cys Gly Gly Gly His Arg Ser Trp  
 20 25 30  
 Ala Phe Ser Asp Thr Glu Ala Ala Met  
 35 40  
  
 <210> 85  
 <211> 42  
 <212> PRT  
 <213> Homo sapiens  
  
 <400> 85  
 Ala Phe Ala Tyr Leu Lys Asp Gly Asp Val Met Leu Tyr Arg Ala Leu  
 1 5 10 15  
 Gly Gly Cys Thr Arg Pro His Val Ile Leu Arg Glu Gly Leu His Gly  
 20 25 30  
 Arg Glu Ile Thr Cys Val Lys Arg Val Gly  
 35 40  
  
 <210> 86  
 <211> 43  
 <212> PRT  
 <213> Homo sapiens  
  
 <400> 86  
 Thr Ile Thr Leu Gly Pro Glu Tyr Gly Val Pro Ser Phe Met Gln Pro  
 1 5 10 15  
 Asp Asp Leu Glu Pro Gly Ser Glu Gly Pro Asp Leu Thr Asp Ile Val  
 20 25 30  
 Ile Thr Cys Ser Glu Asp Thr Thr Val Cys Val  
 35 40  
  
 <210> 87

<211> 41  
<212> PRT  
<213> Homo sapiens

<400> 87  
Leu Ala Leu Pro Thr Thr Gly Ser Ala His Ala Leu Thr Ala Val  
1 5 10 15  
Cys Asn His Ile Ser Ser Val Arg Ala Val Ala Val Trp Gly Ile Gly  
20 25 30  
Thr Pro Gly Gly Pro Gln Asp Pro Gln  
35 40

<210> 88  
<211> 40  
<212> PRT  
<213> Homo sapiens

<400> 88  
Pro Gly Leu Thr Ala His Val Val Ser Ala Gly Gly Arg Ala Glu Met  
1 5 10 15  
His Cys Phe Ser Ile Met Val Thr Pro Asp Pro Ser Thr Pro Ser Arg  
20 25 30  
Leu Ala Cys His Val Met His Leu  
35 40

<210> 89  
<211> 26  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (1)  
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 89  
Xaa Ser His Arg Leu Asp Glu Tyr Trp Asp Arg Gln Arg Asn Arg His  
1 5 10 15  
Arg Met Val Lys Val Asp Pro Glu Thr Arg  
20 25

<210> 90  
<211> 88  
<212> PRT  
<213> Homo sapiens

<400> 90

55

Leu Met Ser Leu Leu Thr Ser Pro His Gln Pro Pro Pro Pro Pro Pro  
 1 5 10 15  
 Ala Ser Ala Ser Pro Ser Ala Val Pro Asn Gly Pro Gln Ser Pro Lys  
 20 25 30  
 Gln Gln Lys Glu Pro Leu Ser His Arg Phe Asn Glu Phe Met Thr Ser  
 35 40 45  
 Lys Pro Lys Ile His Cys Phe Arg Ser Leu Lys Arg Gly Val Ser Ser  
 50 55 60  
 Ala Pro Glu Ser Cys Leu Ser Gly Val Leu Trp Leu His Val Trp Phe  
 65 70 75 80  
 Cys Ile Thr Asn Phe Val Cys Glu  
 85

<210> 91  
 <211> 53  
 <212> PRT  
 <213> Homo sapiens

<400> 91  
 Phe Gln Asn Ala Lys Glu Glu Ala Ser Val Leu Pro Tyr Val Glu Thr  
 1 5 10 15  
 Val Phe Leu Phe Gly Gly Gly Ile Phe Ala Met Ala Leu Cys Leu Ile  
 20 25 30  
 Ser Asp Ala Leu Ser Ser Tyr Arg Asp Ser His Thr Asn Arg Val Leu  
 35 40 45  
 Thr Ser Pro Pro Phe  
 50

<210> 92  
 <211> 45  
 <212> PRT  
 <213> Homo sapiens

<400> 92  
 Arg Leu Met Pro Phe Pro Pro Ser Ser Pro Arg Leu Leu Val Thr Leu  
 1 5 10 15  
 Ala Gly Arg Glu Asp Val Val Gly His Ser Cys Asn Thr Leu Ser Ala  
 20 25 30  
 His Leu Leu Glu Ile Val Thr Met Leu Ile Thr Trp Phe  
 35 40 45

<210> 93

```

<400> 93
Gly Gly Xaa Asp Asp Asp Glu Gly Pro Tyr Thr Pro Phe Asp Thr Pro
 1             5             10             15

Ser Gly Lys Leu Glu Thr Val Lys Trp Ala Phe Thr Trp Pro Leu Ser
      20             25             30

Phe Val Leu Tyr Phe Thr Val Pro Asn Cys Asn Lys Pro Arg Trp Glu
      35             40             45

Lys Trp Phe
      50

```

```
<220>
<221> SITE
<222> (99)
<223> Xaa equals any of the naturally occurring L-amino acids
```

```

<400> 94
Gly Gly Pro Arg Met Lys Arg Ser Gly Asn Pro Gly Ala Glu Val Thr
 1             5             10             15
Asn Ser Ser Val Ala Gly Pro Asp Cys Cys Gly Gly Leu Gly Asn Ile
      20             25             30
Asp Phe Arg Gln Ala Asp Phe Cys Val Met Thr Arg Leu Leu Gly Tyr
      35             40             45
Val Asp Pro Leu Asp Pro Ser Phe Val Ala Ala Val Ile Thr Ile Thr
      50             55             60
Phe Asn Pro Leu Tyr Trp Asn Val Val Ala Arg Trp Glu His Lys Thr
      65             70             75             80
Arg Lys Leu Ser Arg Ala Phe Gly Ser Pro Tyr Leu Ala Cys Tyr Ser
      85             90             95
Leu Ser Xaa Thr Ile Leu Leu Leu Asn Phe Leu Arg Ser His Cys Phe
      100             105             110

```

Thr Gln Ala  
115

<210> 95  
<211> 51  
<212> PRT  
<213> Homo sapiens

<400> 95  
Gly Gly Pro Arg Met Lys Arg Ser Gly Asn Pro Gly Ala Glu Val Thr  
1 5 10 15

Asn Ser Ser Val Ala Gly Pro Asp Cys Cys Gly Gly Leu Gly Asn Ile  
20 25 30

Asp Phe Arg Gln Ala Asp Phe Cys Val Met Thr Arg Leu Leu Gly Tyr  
35 40 45

Val Asp Pro  
50

<210> 96  
<211> 64  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (48)  
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 96  
Leu Asp Pro Ser Phe Val Ala Ala Val Ile Thr Ile Thr Phe Asn Pro  
1 5 10 15

Leu Tyr Trp Asn Val Val Ala Arg Trp Glu His Lys Thr Arg Lys Leu  
20 25 30

Ser Arg Ala Phe Gly Ser Pro Tyr Leu Ala Cys Tyr Ser Leu Ser Xaa  
35 40 45

Thr Ile Leu Leu Leu Asn Phe Leu Arg Ser His Cys Phe Thr Gln Ala  
50 55 60

<210> 97  
<211> 253  
<212> PRT  
<213> Homo sapiens

&lt;400&gt; 97

Pro Gln Arg Ser Glu Leu Ala Ala Ala Ser Asn Arg Pro Cys Arg Val  
 1 5 10 15  
 Cys Ile Ser Leu Leu Leu Cys Leu Glu Asp Arg Thr Met Pro Lys Lys  
 20 25 30  
 Ala Lys Pro Thr Gly Ser Gly Lys Glu Glu Gly Pro Ala Pro Cys Lys  
 35 40 45  
 Gln Met Lys Leu Glu Ala Ala Gly Gly Pro Ser Ala Leu Asn Phe Asp  
 50 55 60  
 Ser Pro Ser Ser Leu Phe Glu Ser Leu Ile Ser Pro Ile Lys Thr Glu  
 65 70 75 80  
 Thr Phe Phe Lys Glu Phe Trp Glu Gln Lys Pro Leu Leu Ile Gln Arg  
 85 90 95  
 Asp Asp Pro Ala Leu Ala Thr Tyr Tyr Gly Ser Leu Phe Lys Leu Thr  
 100 105 110  
 Asp Leu Lys Ser Leu Cys Ser Arg Gly Met Tyr Tyr Gly Arg Asp Val  
 115 120 125  
 Asn Val Cys Arg Cys Val Asn Gly Lys Lys Lys Val Leu Asn Lys Asp  
 130 135 140  
 Gly Lys Ala His Phe Leu Gln Leu Arg Lys Asp Phe Asp Gln Lys Arg  
 145 150 155 160  
 Ala Thr Ile Gln Phe His Gln Pro Gln Arg Phe Lys Asp Glu Leu Trp  
 165 170 175  
 Arg Ile Gln Glu Lys Leu Glu Cys Tyr Phe Gly Ser Leu Val Gly Ser  
 180 185 190  
 Asn Val Tyr Ile Thr Pro Ala Asp Leu Arg Ala Cys Arg Pro Ile Met  
 195 200 205  
 Met Met Ser Arg Phe Ser Ser Cys Ser Trp Arg Glu Arg Asn Thr Gly  
 210 215 220  
 Ala Ser Thr Thr Pro Leu Cys Pro Trp His Glu Ser Thr Ala Trp Arg  
 225 230 235 240  
 Pro Arg Lys Gly Ser Ala Gly Arg Cys Met Ser Leu Cys  
 245 250

&lt;210&gt; 98

&lt;211&gt; 44

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 98

Pro Gln Arg Ser Glu Leu Ala Ala Ala Ser Asn Arg Pro Cys Arg Val  
1 5 10 15

Cys Ile Ser Leu Leu Cys Leu Glu Asp Arg Thr Met Pro Lys Lys  
20 25 30

Ala Lys Pro Thr Gly Ser Gly Lys Glu Glu Gly Pro  
35 40

&lt;210&gt; 99

&lt;211&gt; 45

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 99

Ala Pro Cys Lys Gln Met Lys Leu Glu Ala Ala Gly Gly Pro Ser Ala  
1 5 10 15

Leu Asn Phe Asp Ser Pro Ser Ser Leu Phe Glu Ser Leu Ile Ser Pro  
20 25 30

Ile Lys Thr Glu Thr Phe Phe Lys Glu Phe Trp Glu Gln  
35 40 45

&lt;210&gt; 100

&lt;211&gt; 44

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 100

Lys Pro Leu Leu Ile Gln Arg Asp Asp Pro Ala Leu Ala Thr Tyr Tyr  
1 5 10 15

Gly Ser Leu Phe Lys Leu Thr Asp Leu Lys Ser Leu Cys Ser Arg Gly  
20 25 30

Met Tyr Tyr Gly Arg Asp Val Asn Val Cys Arg Cys  
35 40

&lt;210&gt; 101

&lt;211&gt; 45

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 101

Val Asn Gly Lys Lys Lys Val Leu Asn Lys Asp Gly Lys Ala His Phe  
1 5 10 15

Leu Gln Leu Arg Lys Asp Phe Asp Gln Lys Arg Ala Thr Ile Gln Phe  
20 25 30



His Gln Pro Gln Arg Phe Lys Asp Glu Leu Trp Arg Ile  
                   35                  40                  45

<210> 102  
 <211> 44  
 <212> PRT  
 <213> Homo sapiens

<400> 102  
 Gln Glu Lys Leu Glu Cys Tyr Phe Gly Ser Leu Val Gly Ser Asn Val  
   1                  5                  10                  15

Tyr Ile Thr Pro Ala Asp Leu Arg Ala Cys Arg Pro Ile Met Met Met  
                   20                  25                  30

Ser Arg Phe Ser Ser Cys Ser Trp Arg Glu Arg Asn  
                   35                  40

<210> 103  
 <211> 31  
 <212> PRT  
 <213> Homo sapiens

<400> 103  
 Thr Gly Ala Ser Thr Thr Pro Leu Cys Pro Trp His Glu Ser Thr Ala  
   1                  5                  10                  15

Trp Arg Pro Arg Lys Gly Ser Ala Gly Arg Cys Met Ser Leu Cys  
                   20                  25                  30

<210> 104  
 <211> 53  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (53)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 104  
 Gly Gly Gly Ile His Arg Leu His Asn Gly Ala Leu Gln Leu Arg Val  
   1                  5                  10                  15

Leu Gln Arg Val Glu His Leu His Leu Leu His His Ala Val Lys His  
                   20                  25                  30

Ile Cys Thr Ala Ser Leu Pro Val Leu His Gly Phe Ile Ala Ala Gln  
                   35                  40                  45

Cys Arg Pro Gly Xaa  
                   50

&lt;210&gt; 105

&lt;211&gt; 151

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 105

Trp Asp Arg Trp Ser Asp Ser Ala Leu Arg Arg Leu Arg Gly Ser Gly  
 1 5 10 15

Asp Leu Ala Gly Glu Leu Glu Glu Leu Glu Glu Glu Arg Ala Ala Cys  
 20 25 30

Gln Gly Cys Arg Ala Arg Arg Pro Trp Glu Leu Phe Gln His Arg Ala  
 35 40 45

Leu Arg Arg Gln Val Thr Ser Leu Val Val Leu Gly Ser Ala Met Glu  
 50 55 60

Leu Cys Gly Asn Asp Ser Val Tyr Ala Tyr Ala Ser Ser Val Phe Arg  
 65 70 75 80

Lys Ala Gly Val Pro Glu Ala Lys Ile Gln Tyr Ala Ile Ile Gly Thr  
 85 90 95

Gly Ser Cys Glu Leu Leu Thr Ala Val Val Ser Val Ser Leu Glu Gly  
 100 105 110

Ala Leu Pro Pro Pro Ala Leu Trp Gly Gly Thr Pro Arg Ser Ser Ala  
 115 120 125

Leu Asn Gln Phe Thr Leu Gln Lys Lys Lys Lys Lys Lys Lys Lys Lys  
 130 135 140

Lys Lys Lys Lys Lys Lys Lys  
 145 150

&lt;210&gt; 106

&lt;211&gt; 37

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 106

Arg Arg Leu Arg Gly Ser Gly Asp Leu Ala Gly Glu Leu Glu Glu Leu  
 1 5 10 15

Glu Glu Glu Arg Ala Ala Cys Gln Gly Cys Arg Ala Arg Arg Pro Trp  
 20 25 30

Glu Leu Phe Gln His  
 35

<210> 107  
<211> 29  
<212> PRT  
<213> Homo sapiens

<400> 107  
Arg Gln Val Thr Ser Leu Val Val Leu Gly Ser Ala Met Glu Leu Cys  
1 5 10 15  
Gly Asn Asp Ser Val Tyr Ala Tyr Ala Ser Ser Val Phe  
20 25

<210> 108  
<211> 34  
<212> PRT  
<213> Homo sapiens

<400> 108  
Thr Gly Ser Cys Glu Leu Leu Thr Ala Val Val Ser Val Ser Leu Glu  
1 5 10 15  
Gly Ala Leu Pro Pro Pro Ala Leu Trp Gly Gly Thr Pro Arg Ser Ser  
20 25 30

Ala Leu

<210> 109  
<211> 49  
<212> PRT  
<213> Homo sapiens

<400> 109  
His Glu Leu Arg Leu Arg Lys Asn Thr Val Lys Phe Ser Leu Tyr Arg  
1 5 10 15  
His Phe Lys Asn Thr Leu Ile Phe Ala Val Leu Ala Ser Ile Val Phe  
20 25 30  
Met Gly Trp Thr Thr Lys Thr Phe Arg Ile Ala Lys Cys Gln Ser Asp  
35 40 45

Trp

<210> 110  
<211> 24  
<212> PRT  
<213> Homo sapiens

<400> 110  
Trp Ile Pro Arg Ala Ala Gly Ile Arg His Glu Glu Ser Ile Ala Gln

63

1                      5                      10                      15  
 Arg Ser Tyr Phe Arg Thr Leu Leu  
                                 20  
  
 <210> 111  
 <211> 104  
 <212> PRT  
 <213> Homo sapiens  
  
 <400> 111  
 Ala Asp Thr Asn Phe Thr Gln Glu Thr Ala Met Thr Met Ile Thr Pro  
   1                      5                      10                      15  
 Ser Ser Lys Leu Thr Leu Thr Lys Gly Asn Lys Ser Trp Ser Ser Thr  
                                 20                      25                      30  
 Ala Val Ala Ala Ala Leu Glu Leu Val Asp Pro Pro Gly Cys Arg Asn  
                                 35                      40                      45  
 Ser Ala Arg Gly Ile Asn Cys Ser Ala Phe Leu Leu Pro Tyr Ser Ser  
                                 50                      55                      60  
 His Val Trp Val Pro Leu Ser Gly Val Val Pro Leu Cys Gln Arg Asn  
   65                      70                      75                      80  
 Gln Gly His Thr Val Trp Val Gln Ile Ile Tyr Ser Arg Ser Ser Phe  
                                 85                      90                      95  
 Thr Asp Val Phe Ile Ser Thr Arg  
                                 100

<210> 112  
 <211> 26  
 <212> PRT  
 <213> Homo sapiens

<400> 112  
 Met Thr Met Ile Thr Pro Ser Ser Lys Leu Thr Leu Thr Lys Gly Asn  
   1                      5                      10                      15  
 Lys Ser Trp Ser Ser Thr Ala Val Ala Ala  
                                 20                      25

<210> 113  
 <211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 113  
 Arg Gly Ile Asn Cys Ser Ala Phe Leu Leu Pro Tyr Ser Ser His Val  
   1                      5                      10                      15

Trp Val Pro Leu  
20

<210> 114  
<211> 24  
<212> PRT  
<213> Homo sapiens

<400> 114  
Val Val Pro Leu Cys Gln Arg Asn Gln Gly His Thr Val Trp Val Gln  
1 5 10 15

Ile Ile Tyr Ser Arg Ser Ser Phe  
20

<210> 115  
<211> 26  
<212> PRT  
<213> Homo sapiens

<400> 115  
Asn Phe Asp Ile Lys Val Leu Asn Ala Gln Arg Ala Gly Tyr Lys Ala  
1 5 10 15

Ala Ile Val His Asn Val Asp Ser Asp Asp  
20 25

<210> 116  
<211> 28  
<212> PRT  
<213> Homo sapiens

<400> 116  
Val Leu Lys Lys Ile Asp Ile Pro Ser Val Phe Ile Gly Glu Ser Ser  
1 5 10 15

Ala Asn Ser Leu Lys Asp Glu Phe Thr Tyr Glu Lys  
20 25

<210> 117  
<211> 30  
<212> PRT  
<213> Homo sapiens

<400> 117  
Pro Glu Phe Ser Leu Pro Leu Glu Tyr Tyr Leu Ile Pro Phe Leu Ile  
1 5 10 15

Ile Val Gly Ile Cys Leu Ile Leu Ile Val Ile Phe Met Ile  
20 25 30

<210> 118  
 <211> 34  
 <212> PRT  
 <213> Homo sapiens

<400> 118  
 Thr Lys Phe Val Gln Asp Arg His Arg Ala Arg Arg Asn Arg Leu Arg  
     1                    5                    10                    15

Lys Asp Gln Leu Lys Lys Leu Pro Val His Lys Phe Lys Lys Gly Asp  
                     20                    25                    30

Glu Tyr

<210> 119  
 <211> 27  
 <212> PRT  
 <213> Homo sapiens

<400> 119  
 Glu Asp Gly Asp Lys Leu Arg Ile Leu Pro Cys Ser His Ala Tyr His  
     1                    5                    10                    15

Cys Lys Cys Val Asp Pro Trp Leu Thr Lys Thr  
                     20                    25

<210> 120  
 <211> 24  
 <212> PRT  
 <213> Homo sapiens

<400> 120  
 Val Val Pro Ser Gln Gly Asp Ser Asp Ser Asp Thr Asp Ser Ser Gln  
     1                    5                    10                    15

Glu Glu Asn Glu Val Thr Glu His  
                     20

<210> 121  
 <211> 29  
 <212> PRT  
 <213> Homo sapiens

<400> 121  
 Gln Ser Phe Gly Ala Leu Ser Glu Ser Arg Ser His Gln Asn Met Thr  
     1                    5                    10                    15

Glu Ser Ser Asp Tyr Glu Glu Asp Asp Asn Glu Asp Thr  
                     20                    25

&lt;210&gt; 122

&lt;211&gt; 259

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 122

Ile	Arg	Arg	Leu	Asp	Cys	Asn	Phe	Asp	Ile	Lys	Val	Leu	Asn	Ala	Gln
1				5					10					15	

Arg	Ala	Gly	Tyr	Lys	Ala	Ala	Ile	Val	His	Asn	Val	Asp	Ser	Asp	Asp
			20					25					30		

Leu	Ile	Ser	Met	Gly	Ser	Asn	Asp	Ile	Glu	Val	Leu	Lys	Lys	Ile	Asp
		35					40					45			

Ile	Pro	Ser	Val	Phe	Ile	Gly	Glu	Ser	Ser	Ala	Asn	Ser	Leu	Lys	Asp
	50					55					60				

Glu	Phe	Thr	Tyr	Glu	Lys	Gly	Gly	His	Leu	Ile	Leu	Val	Pro	Glu	Phe
65					70					75					80

Ser	Leu	Pro	Leu	Glu	Tyr	Tyr	Leu	Ile	Pro	Phe	Leu	Ile	Ile	Val	Gly
				85					90					95	

Ile	Cys	Leu	Ile	Leu	Ile	Val	Ile	Phe	Met	Ile	Thr	Lys	Phe	Val	Gln
		100						105					110		

Asp	Arg	His	Arg	Ala	Arg	Arg	Asn	Arg	Leu	Arg	Lys	Asp	Gln	Leu	Lys
		115					120					125			

Lys	Leu	Pro	Val	His	Lys	Phe	Lys	Lys	Gly	Asp	Glu	Tyr	Asp	Val	Cys
	130					135					140				

Ala	Ile	Cys	Leu	Asp	Glu	Tyr	Glu	Asp	Gly	Asp	Lys	Leu	Arg	Ile	Leu
145					150					155					160

Pro	Cys	Ser	His	Ala	Tyr	His	Cys	Lys	Cys	Val	Asp	Pro	Trp	Leu	Thr
				165					170					175	

Lys	Thr	Lys	Lys	Thr	Cys	Pro	Val	Cys	Lys	Gln	Lys	Val	Val	Pro	Ser
		180						185					190		

Gln	Gly	Asp	Ser	Asp	Ser	Asp	Thr	Asp	Ser	Ser	Gln	Glu	Glu	Asn	Glu
		195					200					205			

Val	Thr	Glu	His	Thr	Pro	Leu	Leu	Arg	Pro	Leu	Ala	Ser	Val	Ser	Ala
	210					215					220				

Gln	Ser	Phe	Gly	Ala	Leu	Ser	Glu	Ser	Arg	Ser	His	Gln	Asn	Met	Thr
225					230					235					240

Glu	Ser	Ser	Asp	Tyr	Glu	Glu	Asp	Asp	Asn	Glu	Asp	Thr	Asp	Ser	Ser
				245					250					255	

Asp Ala Glu

<210> 123  
 <211> 36  
 <212> PRT  
 <213> Homo sapiens

<400> 123  
 Ala Gln Cys Ser Ile Tyr Leu Ile Gln Val Ile Phe Gly Ala Val Asp  
     1                    5                    10                    15  
 Leu Pro Ala Lys Leu Val Gly Phe Leu Val Ile Asn Ser Leu Gly Arg  
                     20                    25                    30  
 Arg Pro Ala Gln  
                     35

<210> 124  
 <211> 188  
 <212> PRT  
 <213> Homo sapiens

<400> 124  
 Gly Thr Val Gln His Leu Pro Asn Pro Gly Asp Leu Trp Cys Cys Gly  
     1                    5                    10                    15  
 Pro Ala Cys Gln Ala Cys Gly Leu Pro Cys His Gln Leu Pro Gly Ser  
                     20                    25                    30  
 Pro Ala Cys Pro Asp Gly Cys Thr Ala Ala Gly Arg His Leu His Pro  
                     35                    40                    45  
 Ala Gln Trp Gly Asp Thr Pro Gly Pro Val His Cys Pro Asn Leu Ser  
     50                    55                    60  
 Cys Cys Ala Gly Glu Gly Leu Ser Gly Cys Leu Leu Gln Leu His Leu  
     65                    70                    75                    80  
 Pro Val Tyr Trp Glu Leu Tyr Pro Thr Met Ile Arg Gln Thr Gly Met  
                     85                    90                    95  
 Gly Met Gly Ser Thr Met Ala Arg Val Gly Ser Ile Val Ser Pro Leu  
                     100                    105                    110  
 Val Ser Met Thr Ala Glu Leu Tyr Pro Ser Met Pro Leu Phe Ile Tyr  
                     115                    120                    125  
 Gly Ala Val Pro Val Ala Ala Ser Ala Val Thr Val Leu Leu Pro Glu  
     130                    135                    140  
 Thr Leu Gly Gln Pro Leu Pro Asp Thr Val Gln Asp Leu Glu Ser Arg



```
<210> 125
<211> 23
<212> PRT
<213> Homo sapiens
```

Cys Gly Leu Pro Cys His Gln  
20

```
<210> 126
<211> 26
<212> PRT
<213> Homo sapiens
```

Thr Pro Gly Pro Val His Cys Pro Asn Leu  
20 25

```
<210> 127
<211> 22
<212> PRT
<213> Homo sapiens
```

<400> 127  
Leu His Leu Pro Val Tyr Trp Glu Leu Tyr Pro Thr Met Ile Arg Gln  
1 5 10 15

Thr Gly Met Gly Met Gly  
20

```
<210> 128
<211> 23
<212> PRT
<213> Homo sapiens
```

<400> 128  
Leu Val Ser Met Thr Ala Glu Leu Tyr Pro Ser Met Pro Leu Phe Ile

1                      5                      10                      15  
 Tyr Gly Ala Val Pro Val Ala  
                                  20  
  
 <210> 129  
 <211> 27  
 <212> PRT  
 <213> Homo sapiens  
  
 <400> 129  
 Pro Asp Thr Val Gln Asp Leu Glu Ser Arg Lys Gly Lys Gln Thr Arg  
   1                      5                      10                      15  
  
 Gln Gln Gln Glu His Gln Lys Tyr Met Val Pro  
                                  20                      25  
  
 <210> 130  
 <211> 720  
 <212> PRT  
 <213> Homo sapiens  
  
 <400> 130  
 Cys Leu Glu Ala Ala Met Ile Glu Gly Glu Ile Glu Ser Leu His Ser  
   1                      5                      10                      15  
  
 Glu Asn Ser Gly Lys Ser Gly Gln Glu His Trp Phe Thr Glu Leu Pro  
                                  20                      25                      30  
  
 Pro Val Leu Thr Phe Glu Leu Ser Arg Phe Glu Phe Asn Gln Ala Leu  
                                  35                      40                      45  
  
 Gly Arg Pro Glu Lys Ile His Asn Lys Leu Glu Phe Pro Gln Val Leu  
                                  50                      55                      60  
  
 Tyr Leu Asp Arg Tyr Met His Arg Asn Arg Glu Ile Thr Arg Ile Lys  
   65                      70                      75                      80  
  
 Arg Glu Glu Ile Lys Arg Leu Lys Asp Tyr Leu Thr Val Leu Gln Gln  
                                  85                      90                      95  
  
 Arg Leu Glu Arg Tyr Leu Ser Tyr Gly Ser Gly Pro Lys Arg Phe Pro  
                                  100                      105                      110  
  
 Leu Val Asp Val Leu Gln Tyr Ala Leu Glu Phe Ala Ser Ser Lys Pro  
                                  115                      120                      125  
  
 Val Cys Thr Ser Pro Val Asp Asp Ile Asp Ala Ser Ser Pro Pro Ser  
   130                      135                      140  
  
 Gly Ser Ile Pro Ser Gln Thr Leu Pro Ser Thr Thr Glu Gln Gln Gly  
   145                      150                      155                      160

Ala Leu Ser Ser Glu Leu Pro Ser Thr Ser Pro Ser Ser Val Ala Ala  
 165 170 175  
 Ile Ser Ser Arg Ser Val Ile His Lys Pro Phe Thr Gln Ser Arg Ile  
 180 185 190  
 Pro Pro Asp Leu Pro Met His Pro Ala Pro Arg His Ile Thr Glu Glu  
 195 200 205  
 Glu Leu Ser Val Leu Glu Ser Cys Leu His Arg Trp Arg Thr Glu Ile  
 210 215 220  
 Glu Asn Asp Thr Arg Asp Leu Gln Glu Ser Ile Ser Arg Ile His Arg  
 225 230 235 240  
 Thr Ile Glu Leu Met Tyr Ser Asp Lys Ser Met Ile Gln Val Pro Tyr  
 245 250 255  
 Arg Leu His Ala Val Leu Val His Glu Gly Gln Ala Asn Ala Gly His  
 260 265 270  
 Tyr Trp Ala Tyr Ile Phe Asp His Arg Glu Ser Arg Trp Met Lys Tyr  
 275 280 285  
 Asn Asp Ile Ala Val Thr Lys Ser Ser Trp Glu Glu Leu Val Arg Asp  
 290 295 300  
 Ser Phe Gly Gly Tyr Arg Asn Ala Ser Ala Tyr Cys Leu Met Tyr Ile  
 305 310 315 320  
 Asn Asp Lys Ala Gln Phe Leu Ile Gln Glu Glu Phe Asn Lys Glu Thr  
 325 330 335  
 Gly Gln Pro Leu Val Gly Ile Glu Thr Leu Pro Pro Asp Leu Arg Asp  
 340 345 350  
 Phe Val Glu Glu Asp Asn Gln Arg Phe Glu Lys Glu Leu Glu Glu Trp  
 355 360 365  
 Asp Ala Gln Leu Ala Gln Lys Ala Leu Gln Glu Lys Leu Leu Ala Ser  
 370 375 380  
 Gln Lys Leu Arg Glu Ser Glu Thr Ser Val Thr Thr Ala Gln Ala Ala  
 385 390 395 400  
 Gly Asp Pro Glu Tyr Leu Glu Gln Pro Ser Arg Ser Asp Phe Ser Lys  
 405 410 415  
 His Leu Lys Glu Glu Thr Ile Gln Ile Ile Thr Lys Ala Ser His Glu  
 420 425 430  
 His Glu Asp Lys Ser Pro Glu Thr Val Leu Gln Ser Ala Ile Lys Leu  
 435 440 445  
 Glu Tyr Ala Arg Leu Val Lys Leu Ala Gln Glu Asp Thr Pro Pro Glu

450		455		460											
Thr	Asp	Tyr	Arg	Leu	His	His	Val	Val	Val	Tyr	Phe	Ile	Gln	Asn	Gln
465					470					475					480
Ala	Pro	Lys	Lys	Ile	Ile	Glu	Lys	Thr	Leu	Leu	Glu	Gln	Phe	Gly	Asp
				485					490					495	
Arg	Asn	Leu	Ser	Phe	Asp	Glu	Arg	Cys	His	Asn	Ile	Met	Lys	Val	Ala
			500					505					510		
Gln	Ala	Lys	Leu	Glu	Met	Ile	Lys	Pro	Glu	Glu	Val	Asn	Leu	Glu	Glu
		515					520					525			
Tyr	Glu	Glu	Trp	His	Gln	Asp	Tyr	Arg	Lys	Phe	Arg	Glu	Thr	Thr	Met
	530					535					540				
Tyr	Leu	Ile	Ile	Gly	Leu	Glu	Asn	Phe	Gln	Arg	Glu	Ser	Tyr	Ile	Asp
545					550				555						560
Ser	Leu	Leu	Phe	Leu	Ile	Cys	Ala	Tyr	Gln	Asn	Asn	Lys	Glu	Leu	Leu
			565						570					575	
Ser	Lys	Gly	Leu	Tyr	Arg	Gly	His	Asp	Glu	Glu	Leu	Ile	Ser	His	Tyr
			580					585					590		
Arg	Arg	Glu	Cys	Leu	Leu	Lys	Leu	Asn	Glu	Gln	Ala	Ala	Glu	Leu	Phe
		595					600						605		
Glu	Ser	Gly	Glu	Asp	Arg	Glu	Val	Asn	Asn	Gly	Leu	Ile	Ile	Met	Asn
	610					615					620				
Glu	Phe	Ile	Val	Pro	Phe	Leu	Pro	Leu	Leu	Leu	Val	Asp	Glu	Met	Glu
625					630					635					640
Glu	Lys	Asp	Ile	Leu	Ala	Val	Glu	Asp	Met	Arg	Asn	Arg	Trp	Cys	Ser
			645						650					655	
Tyr	Leu	Gly	Gln	Glu	Met	Glu	Pro	His	Leu	Gln	Glu	Lys	Leu	Thr	Asp
		660						665						670	
Phe	Leu	Pro	Lys	Leu	Leu	Asp	Cys	Ser	Met	Glu	Ile	Lys	Ser	Phe	His
	675						680					685			
Glu	Pro	Pro	Lys	Leu	Pro	Ser	Tyr	Ser	Thr	His	Glu	Leu	Cys	Glu	Arg
	690					695					700				
Phe	Ala	Arg	Ile	Met	Leu	Ser	Leu	Ser	Arg	Thr	Pro	Ala	Asp	Gly	Arg
705					710					715					720

&lt;211&gt; 24

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 131

Met Ile Glu Gly Glu Ile Glu Ser Leu His Ser Glu Asn Ser Gly Lys  
1 5 10 15

Ser Gly Gln Glu His Trp Phe Thr  
20

&lt;210&gt; 132

&lt;211&gt; 25

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 132

Phe Glu Leu Ser Arg Phe Glu Phe Asn Gln Ala Leu Gly Arg Pro Glu  
1 5 10 15

Lys Ile His Asn Lys Leu Glu Phe Pro  
20 25

&lt;210&gt; 133

&lt;211&gt; 26

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 133

Glu Ile Thr Arg Ile Lys Arg Glu Glu Ile Lys Arg Leu Lys Asp Tyr  
1 5 10 15

Leu Thr Val Leu Gln Gln Arg Leu Glu Arg  
20 25

&lt;210&gt; 134

&lt;211&gt; 27

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 134

Pro Lys Arg Phe Pro Leu Val Asp Val Leu Gln Tyr Ala Leu Glu Phe  
1 5 10 15

Ala Ser Ser Lys Pro Val Cys Thr Ser Pro Val  
20 25

&lt;210&gt; 135

&lt;211&gt; 26

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 135

Ile Pro Ser Gln Thr Leu Pro Ser Thr Thr Glu Gln Gln Gly Ala Leu  
1 5 10 15

Ser Ser Glu Leu Pro Ser Thr Ser Pro Ser  
20 25

&lt;210&gt; 136

&lt;211&gt; 24

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 136

Ser Val Ile His Lys Pro Phe Thr Gln Ser Arg Ile Pro Pro Asp Leu  
1 5 10 15

Pro Met His Pro Ala Pro Arg His  
20

&lt;210&gt; 137

&lt;211&gt; 23

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 137

Cys Leu His Arg Trp Arg Thr Glu Ile Glu Asn Asp Thr Arg Asp Leu  
1 5 10 15

Gln Glu Ser Ile Ser Arg Ile  
20

&lt;210&gt; 138

&lt;211&gt; 28

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 138

Lys Ser Met Ile Gln Val Pro Tyr Arg Leu His Ala Val Leu Val His  
1 5 10 15

Glu Gly Gln Ala Asn Ala Gly His Tyr Trp Ala Tyr  
20 25

&lt;210&gt; 139

&lt;211&gt; 29

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 139

Arg Trp Met Lys Tyr Asn Asp Ile Ala Val Thr Lys Ser Ser Trp Glu

1 5 10 15  
Glu Leu Val Arg Asp Ser Phe Gly Gly Tyr Arg Asn Ala  
20 25

<210> 140  
<211> 24  
<212> PRT  
<213> Homo sapiens

<400> 140  
Ile Asn Asp Lys Ala Gln Phe Leu Ile Gln Glu Glu Phe Asn Lys Glu  
1 5 10 15

Thr Gly Gln Pro Leu Val Gly Ile  
20

<210> 141  
<211> 23  
<212> PRT  
<213> Homo sapiens

<400> 141  
Met Ile Gln Val Pro Tyr Arg Leu His Ala Val Leu Val His Glu Gly  
1 5 10 15

Gln Ala Asn Ala Gly His Tyr  
20

<210> 142  
<211> 26  
<212> PRT  
<213> Homo sapiens

<400> 142  
Asp Asn Gln Arg Phe Glu Lys Glu Leu Glu Glu Trp Asp Ala Gln Leu  
1 5 10 15

Ala Gln Lys Ala Leu Gln Glu Lys Leu Leu  
20 25

<210> 143  
<211> 23  
<212> PRT  
<213> Homo sapiens

<400> 143  
Ser Glu Thr Ser Val Thr Thr Ala Gln Ala Ala Gly Asp Pro Glu Tyr  
1 5 10 15

Leu Glu Gln Pro Ser Arg Ser

20

<210> 144  
<211> 28  
<212> PRT  
<213> Homo sapiens

<400> 144  
Gln Ile Ile Thr Lys Ala Ser His Glu His Glu Asp Lys Ser Pro Glu  
1 5 10 15

Thr Val Leu Gln Ser Ala Ile Lys Leu Glu Tyr Ala  
20 25

<210> 145  
<211> 28  
<212> PRT  
<213> Homo sapiens

<400> 145  
Leu Ala Gln Glu Asp Thr Pro Pro Glu Thr Asp Tyr Arg Leu His His  
1 5 10 15

Val Val Val Tyr Phe Ile Gln Asn Gln Ala Pro Lys  
20 25

<210> 146  
<211> 29  
<212> PRT  
<213> Homo sapiens

<400> 146  
Gly Asp Arg Asn Leu Ser Phe Asp Glu Arg Cys His Asn Ile Met Lys  
1 5 10 15

Val Ala Gln Ala Lys Leu Glu Met Ile Lys Pro Glu Glu  
20 25

<210> 147  
<211> 26  
<212> PRT  
<213> Homo sapiens

<400> 147  
Glu Glu Trp His Gln Asp Tyr Arg Lys Phe Arg Glu Thr Thr Met Tyr  
1 5 10 15

Leu Ile Ile Gly Leu Glu Asn Phe Gln Arg  
20 25



<210> 148  
<211> 29  
<212> PRT  
<213> Homo sapiens

<400> 148  
Ile Cys Ala Tyr Gln Asn Asn Lys Glu Leu Leu Ser Lys Gly Leu Tyr  
1 5 10 15  
Arg Gly His Asp Glu Glu Leu Ile Ser His Tyr Arg Arg  
20 25

<210> 149  
<211> 28  
<212> PRT  
<213> Homo sapiens

<400> 149  
Cys Leu Leu Lys Leu Asn Glu Gln Ala Ala Glu Leu Phe Glu Ser Gly  
1 5 10 15  
Glu Asp Arg Glu Val Asn Asn Gly Leu Ile Ile Met  
20 25

<210> 150  
<211> 31  
<212> PRT  
<213> Homo sapiens

<400> 150  
Val Asp Glu Met Glu Glu Lys Asp Ile Leu Ala Val Glu Asp Met Arg  
1 5 10 15  
Asn Arg Trp Cys Ser Tyr Leu Gly Gln Glu Met Glu Pro His Leu  
20 25 30

<210> 151  
<211> 25  
<212> PRT  
<213> Homo sapiens

<400> 151  
Gln Glu Lys Leu Thr Asp Phe Leu Pro Lys Leu Leu Asp Cys Ser Met  
1 5 10 15  
Glu Ile Lys Ser Phe His Glu Pro Pro  
20 25

<210> 152  
<211> 21  
<212> PRT

<213> Homo sapiens

<400> 152

Gln Ile Ala Thr Ser Val His His Asn Ile Asn Arg Lys Lys Arg Ser  
1 5 10 15

Val Leu Arg Leu Leu  
20

<210> 153

<211> 32

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (10)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (22)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 153

Pro Leu Leu Arg Gly Leu Phe Ile Arg Xaa Arg Ala Gly His Tyr Glu  
1 5 10 15

Cys Val Phe His Glu Xaa Val Glu Gly Gly Ala Cys Cys Glu Gln Cys  
20 25 30

<210> 154

<211> 44

<212> PRT

<213> Homo sapiens

<400> 154

Leu Val Asn Asn Ser Phe Phe Leu Glu Phe Ile Tyr Arg Pro Asp Ser  
1 5 10 15

Lys Asn Trp Gln Tyr Gln Glu Thr Ile Lys Lys Gly Asp Leu Leu Leu  
20 25 30

Asn Arg Val Gln Lys Leu Ser Arg Val Ile Asn Met  
35 40

<210> 155

<211> 34

<212> PRT

<213> Homo sapiens

<400> 155

Ile Arg Glu Leu Ser Arg Phe Ile Ala Ala Gly Arg Leu His Cys Lys  
1 5 10 15

Ile Asp Lys Val Asn Glu Ile Val Glu Thr Asn Arg Tyr Ser His Phe  
20 25 30

Ser Glu

<210> 156

<211> 195

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (11)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 156

Gly Ser Gln Pro Pro Gly Pro Val Pro Glu Xaa Leu Ile Arg Ile Tyr  
1 5 10 15

Ser Met Arg Phe Cys Pro Tyr Ser His Arg Thr Arg Leu Val Leu Lys  
20 25 30

Ala Lys Asp Ile Arg His Glu Val Val Asn Ile Asn Leu Arg Asn Lys  
35 40 45

Pro Glu Trp Tyr Tyr Thr Lys His Pro Phe Gly His Ile Pro Val Leu  
50 55 60

Glu Thr Ser Gln Cys Gln Leu Ile Tyr Glu Ser Val Ile Ala Cys Glu  
65 70 75 80

Tyr Leu Asp Asp Ala Tyr Pro Gly Arg Lys Leu Phe Pro Tyr Asp Pro  
85 90 95

Tyr Glu Arg Ala Arg Gln Lys Met Leu Leu Glu Leu Phe Cys Lys Val  
100 105 110

Pro His Leu Thr Lys Glu Cys Leu Val Ala Leu Arg Cys Gly Arg Glu  
115 120 125

Cys Thr Asn Leu Lys Ala Ala Leu Arg Gln Glu Phe Ser Asn Leu Glu  
130 135 140

Glu Ile Leu Glu Tyr Gln Asn Thr Thr Phe Phe Gly Gly Thr Cys Ile  
145 150 155 160

Ser Met Ile Asp Tyr Leu Leu Trp Pro Trp Phe Glu Arg Leu Asp Val

165 170 175  
Tyr Gly Ile Leu Asp Cys Val Ser His Thr Pro Ala Cys Gly Ser Gly  
180 185 190

Tyr Gln Pro  
195

<210> 157  
<211> 14  
<212> PRT  
<213> Homo sapiens

<400> 157  
Leu Ala Ser Pro Phe Pro Val Pro Leu His Arg Cys Ser Ala  
1 5 10

<210> 158  
<211> 29  
<212> PRT  
<213> Homo sapiens

<400> 158  
Met Arg Phe Cys Pro Tyr Ser His Arg Thr Arg Leu Val Leu Lys Ala  
1 5 10 15

Lys Asp Ile Arg His Glu Val Val Asn Ile Asn Leu Arg  
20 25

<210> 159  
<211> 24  
<212> PRT  
<213> Homo sapiens

<400> 159  
Asn Lys Pro Glu Trp Tyr Tyr Thr Lys His Pro Phe Gly His Ile Pro  
1 5 10 15

Val Leu Glu Thr Ser Gln Cys Gln  
20

<210> 160  
<211> 24  
<212> PRT  
<213> Homo sapiens

<400> 160  
Lys Leu Phe Pro Tyr Asp Pro Tyr Glu Arg Ala Arg Gln Lys Met Leu  
1 5 10 15

Leu Glu Leu Phe Cys Lys Val Pro

20

<210> 161  
 <211> 25  
 <212> PRT  
 <213> Homo sapiens

<400> 161  
 Val Ala Leu Arg Cys Gly Arg Glu Cys Thr Asn Leu Lys Ala Ala Leu  
           1                  5                  10                  15  
 Arg Gln Glu Phe Ser Asn Leu Glu Glu  
                   20                  25

<210> 162  
 <211> 24  
 <212> PRT  
 <213> Homo sapiens

<400> 162  
 Ser Met Ile Asp Tyr Leu Leu Trp Pro Trp Phe Glu Arg Leu Asp Val  
           1                  5                  10                  15  
 Tyr Gly Ile Leu Asp Cys Val Ser  
                   20

<210> 163  
 <211> 60  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (15)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 163  
 Ala Ala Gly Cys Val Trp Asp Thr Gly Leu Cys Glu Pro His Xaa Ser  
           1                  5                  10                  15  
 Leu Arg Leu Trp Ile Ser Ala Met Lys Trp Asp Pro Thr Val Cys Ala  
                   20                  25                  30  
 Leu Leu Met Asp Lys Ser Ile Phe Gln Gly Phe Leu Asn Leu Tyr Phe  
           35                  40                  45  
 Gln Asn Asn Pro Asn Ala Phe Asp Phe Gly Leu Cys  
           50                  55                  60

<210> 164  
 <211> 180

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 164

Val Tyr Leu Phe Leu Thr Tyr Arg Gln Ala Val Val Ile Ala Leu Leu  
 1 5 10 15

Val Lys Val Gly Val Ile Ser Glu Lys His Thr Trp Glu Trp Gln Thr  
 20 25 30

Val Glu Ala Val Ala Thr Gly Leu Gln Asp Phe Ile Ile Cys Ile Glu  
 35 40 45

Met Phe Leu Ala Ala Ile Ala His His Tyr Thr Phe Ser Tyr Lys Pro  
 50 55 60

Tyr Val Gln Glu Ala Glu Glu Gly Ser Cys Phe Asp Ser Phe Leu Ala  
 65 70 75 80

Met Trp Asp Val Ser Asp Ile Arg Asp Asp Ile Ser Glu Gln Val Arg  
 85 90 95

His Val Gly Arg Thr Val Arg Gly His Pro Arg Lys Lys Leu Phe Pro  
 100 105 110

Glu Asp Gln Asp Gln Asn Glu His Thr Ser Leu Leu Ser Ser Ser Ser  
 115 120 125

Gln Asp Ala Ile Ser Ile Ala Ser Ser Met Pro Pro Ser Pro Met Gly  
 130 135 140

His Tyr Gln Gly Phe Gly His Thr Val Thr Pro Gln Thr Thr Pro Thr  
 145 150 155 160

Thr Ala Lys Ile Ser Asp Glu Ile Leu Ser Asp Thr Ile Gly Glu Lys  
 165 170 175

Lys Glu Pro Ser  
 180

&lt;210&gt; 165

&lt;211&gt; 176

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 165

Thr Asn Asn Lys Asp Ser Leu Gly Trp Tyr Leu Phe Thr Val Leu Asp  
 1 5 10 15

Ser Trp Ile Ala Leu Lys Tyr Pro Gly Ile Ala Ile Tyr Val Asp Thr  
 20 25 30

Cys Arg Glu Cys Tyr Glu Ala Tyr Val Ile Tyr Asn Phe Met Gly Phe  
 35 40 45

Leu Thr Asn Tyr Leu Thr Asn Arg Tyr Pro Asn Leu Val Leu Ile Leu  
50 55 60

Glu Ala Lys Asp Gln Gln Lys His Phe Pro Pro Leu Cys Cys Cys Pro  
65 70 75 80

Pro Trp Ala Met Gly Glu Val Leu Leu Phe Arg Cys Lys Leu Ser Val  
85 90 95

Leu Gln Tyr Thr Val Val Arg Pro Phe Thr Thr Ile Val Ala Leu Ile  
100 105 110

Cys Glu Leu Leu Gly Ile Tyr Asp Glu Gly Asn Phe Ser Phe Ser Asn  
115 120 125

Ala Trp Thr Tyr Leu Val Ile Ile Asn Asn Met Ser Gln Leu Phe Ala  
130 135 140

Met Tyr Cys Leu Leu Leu Phe Tyr Lys Val Leu Lys Glu Glu Leu Ser  
145 150 155 160

Pro Ile Gln Pro Val Gly Lys Phe Leu Cys Val Lys Leu Val Val Phe  
165 170 175

<210> 166

<211> 28

<212> PRT

<213> Homo sapiens

<400> 166

Gln Asn Ser Gln Arg Thr Gly Leu Pro Ile Thr Ile Phe Ser Arg Ser  
1 5 10 15

Phe Pro Leu Leu Thr Gly Ser Asp Leu Cys Glu Asn  
20 25

<210> 167

<211> 9

<212> PRT

<213> Homo sapiens

<400> 167

Gln Phe Phe Leu Cys Arg Asp Cys Ser  
1 5

<210> 168

<211> 38

<212> PRT

<213> Homo sapiens

<400> 168

Glu Arg Glu Ser Cys Ser Ile Ile Gln Ala Gly Val Gln Trp Cys Asn  
1 5 10 15

Leu Ser Ser Leu Arg Pro Pro Pro Pro Gly Phe Lys Gln Phe Ser His  
20 25 30

Leu Ser Leu Pro Ser Ser  
35

<210> 169

<211> 116

<212> PRT

<213> Homo sapiens

<400> 169

Leu Arg Glu Asn Leu Ala Leu Ser Ser Arg Leu Glu Cys Ser Gly Ala  
1 5 10 15

Ile Ser Ala His Cys Asp Leu His Leu Leu Gly Ser Ser Asn Ser Pro  
20 25 30

Thr Ser Ala Ser Gln Val Val Arg Thr Thr Gly Ala His His Gln Ala  
35 40 45

Gln Pro Ile Phe Val Phe Leu Val Glu Thr Gly Phe His His Val Gly  
50 55 60

Gln Ala His Leu Lys Gln Leu Thr Ser Arg Tyr Pro Pro His Leu Ala  
65 70 75 80

Ser Gln Ser Ala Gly Ile Thr Gly Met Ser Tyr Arg Thr Gln Pro Lys  
85 90 95

Leu Leu Trp Phe Tyr Leu Tyr Lys Gln Phe Lys Gln Tyr Arg Glu Val  
100 105 110

Gly Ser Arg Lys  
115

<210> 170

<211> 25

<212> PRT

<213> Homo sapiens

<400> 170

Ser Ser Arg Leu Glu Cys Ser Gly Ala Ile Ser Ala His Cys Asp Leu  
1 5 10 15

His Leu Leu Gly Ser Ser Asn Ser Pro  
20 25



<210> 171  
 <211> 40  
 <212> PRT  
 <213> Homo sapiens

<400> 171  
 Gly Ala His His Gln Ala Gln Pro Ile Phe Val Phe Leu Val Glu Thr  
           1                  5                  10                  15  
 Gly Phe His His Val Gly Gln Ala His Leu Lys Gln Leu Thr Ser Arg  
                   20                  25                  30  
 Tyr Pro Pro His Leu Ala Ser Gln  
           35                  40

<210> 172  
 <211> 25  
 <212> PRT  
 <213> Homo sapiens

<400> 172  
 Ile Thr Gly Met Ser Tyr Arg Thr Gln Pro Lys Leu Leu Trp Phe Tyr  
           1                  5                  10                  15  
 Leu Tyr Lys Gln Phe Lys Gln Tyr Arg  
                   20                  25

<210> 173  
 <211> 25  
 <212> PRT  
 <213> Homo sapiens

<400> 173  
 Glu Asn Phe Pro Glu Thr Arg Glu Val Arg Ala Phe Ser Pro Arg Glu  
           1                  5                  10                  15  
 Asn Leu Glu Leu Cys Thr Cys Lys Ser  
                   20                  25

<210> 174  
 <211> 11  
 <212> PRT  
 <213> Homo sapiens

<400> 174  
 Ala Leu Tyr Cys Ser Pro Ser Leu Gln Ile Asp  
           1                  5                  10

<210> 175

&lt;211&gt; 37

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 175

Cys His Cys Ser Met Leu Lys Ser His Gly Asp Val Gln Asn Val Leu  
1 5 10 15

Thr Leu Phe Val Thr Val Leu Ser Asp Val Ser Tyr Leu Gln Gln Ile  
20 25 30

Gln Lys Lys Leu Arg  
35

&lt;210&gt; 176

&lt;211&gt; 39

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 176

Cys Tyr Phe His Gln Lys Ala Gln Ser Asn Gly Pro Glu Lys Gln Glu  
1 5 10 15

Lys Glu Gly Val Ile Gln Asn Phe Lys Arg Thr Leu Ser Lys Lys Glu  
20 25 30

Lys Lys Glu Lys Lys Lys Lys  
35

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A.</b> The indications made below relate to the microorganism referred to in the description on page <u>54</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit  <u>February 12, 1998</u>	Accession Number  <u>209627</u>
<b>C. ADDITIONAL INDICATIONS</b> (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (if the indications are not for all designated States)	
EUROPE In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<p style="text-align: center;"><b>For receiving Office use only</b></p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <hr/> <p>Authorized officer <u>Lydia Meadows</u>  <u>Paralegal Specialist</u>  <u>IAPD-PCT Operations</u>  <u>(703) 306-3745</u></p>	<p style="text-align: center;"><b>For International Bureau use only</b></p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <hr/> <p>Authorized officer</p>
--	---

**CANADA**

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

**NORWAY**

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/05721

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6) :C07K 14/435, 14/47; C12N 1/21, 5/00, 15/12, 15/63  
US CL :435/69.1, 325, 243, 320.1; 536/23.1, 24.3; 530/300, 399

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 325, 243, 320.1; 536/23.1, 24.3; 530/300, 399

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, GENEMBL, GENBANK, MEDLINE, GENESEQ  
search terms: secreted protein

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GenBank Database, Institute for Genomic Research, Rockville, Maryland USA, Accession No. G20793, Adams, M.D., Human STS A006G35, 24 JULY 1996.	1-3, 5-6
X	GenBank Database, Washington University School of Medicine, St. Louis, MO USA, Accession No. AA477432, HILLIER et al., zu42f03.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 740669 3', 09 NOVEMBER 1997.	1-3, 5-6
X	GenBank Database, Washington University School of Medicine, St. Louis, MO USA, Accession No. AA234651, HILLIER et al., zr75f06.rl Soares NhHMPu S1 Homo sapiens cDNA clone 669251 5'.	1-3 and 5-6

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*B* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 28 MAY 1999	Date of mailing of the international search report 17 JUN 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer <i>D. Lawrence For</i> CHRISTINE SAOUD Telephone No. (703) 308-0196

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/05721

**Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-12,14-16,21 with regard to SEQ ID NO:11 and 44

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/05721

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-12,14-16,21 with regard to SEQ ID NO:11 and 44

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/05721

### BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-12, 14-16, and 21, drawn to polynucleotides, polypeptides, vectors, host cells, and methods of making the polypeptides.

Group II, claim(s) 13, drawn to antibodies to the polypeptides.

Group III, claim(s) 17, drawn to a method of treatment by administration of the polypeptide or polynucleotide.

Group IV, claim 18, drawn to a method of diagnosing a condition by measuring for the polynucleotide.

Group V, claim 19, drawn to a method of diagnosing a condition by measuring for the polypeptide.

Group VI, claims 20 and 23, drawn to a method of identifying a binding partner to the polypeptide.

Group VII, claim 22, drawn to a method of identifying an activity for a protein.

The inventions listed as Groups I-VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The main invention is Group I which is first product, first method of making the product, and the first method of using the product. Pursuant to 37 CFR 1.475(d), these claims are considered by the ISA/US to constitute the main invention, and none of the related groups II-VII correspond to the main invention. The special technical feature of Group I is the polynucleotide, which is not shared by the other inventive groups.

The product of Group II does not share the same or corresponding technical feature with Group I because the products have materially different structures and functions, and each defines a separate invention over the art. The methods of Groups III-VII do not share the same or corresponding special technical feature with Group I because the methods have materially different process steps and are practiced for materially different purposes, and each defines a separate invention over the art.

Additionally, each of the inventions of Groups I-VII are directed to or use one of 31 distinct polynucleotides or polypeptides which lack unity of invention. Nucleotide sequences encoding different proteins are structurally distinct chemical compounds and are unrelated to one another. They do not share a common structure which provides for a common function, and therefore, they lack unity of invention. The 31 different polynucleotides and corresponding polypeptides are enumerated in Table 1 at pages 54-56 of the description.

**BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING**

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-12, 14-16, and 21, drawn to polynucleotides, polypeptides, vectors, host cells, and methods of making the polypeptides.

Group II, claim(s) 13, drawn to antibodies to the polypeptides.

Group III, claim(s) 17, drawn to a method of treatment by administration of the polypeptide or polynucleotide.

Group IV, claim 18, drawn to a method of diagnosing a condition by measuring for the polynucleotide.

Group V, claim 19, drawn to a method of diagnosing a condition by measuring for the polypeptide.

Group VI, claims 20 and 23, drawn to a method of identifying a binding partner to the polypeptide.

Group VII, claim 22, drawn to a method of identifying an activity for a protein.

The inventions listed as Groups I-VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The main invention is Group I which is first product, first method of making the product, and the first method of using the product. Pursuant to 37 CFR 1.475(d), these claims are considered by the ISA/US to constitute the main invention, and none of the related groups II-VII correspond to the main invention. The special technical feature of Group I is the polynucleotide, which is not shared by the other inventive groups.

The product of Group II does not share the same or corresponding technical feature with Group I because the products have materially different structures and functions, and each defines a separate invention over the art. The methods of Groups III-VII do not share the same or corresponding special technical feature with Group I because the methods have materially different process steps and are practiced for materially different purposes, and each defines a separate invention over the art.

Additionally, each of the inventions of Groups I-VII are directed to or use one of 31 distinct polynucleotides or polypeptides which lack unity of invention. Nucleotide sequences encoding different proteins are structurally distinct chemical compounds and are unrelated to one another. They do not share a common structure which provides for a common function, and therefore, they lack unity of invention. The 31 different polynucleotides and corresponding polypeptides are enumerated in Table 1 at pages 54-56 of the description.